

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:13:22 ; Search time 70.93 Seconds  
(without alignments)  
111.842 Million cell updates/sec

Title: US-09-389-782A-1

Perfect score: 1263

Sequence: 1 EPKSCDKTHTPCPAPPELL.....MHEALNHYTKLSLSPGK 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_36.\*

1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1263	100.0	232	18 W26232	Human IgG1 hinge/F
2	1263	100.0	259	20 Y24154	Protein from pCds1
3	1263	100.0	329	17 R91806	Human immunoglobul
4	1263	100.0	351	14 R43685	Human kappa immuno
5	1263	100.0	371	10 P91918	Sequence of the li
6	1263	100.0	371	10 P93558	Linked human IgG
7	1263	100.0	376	19 W60037	Antigenic peptide
8	1263	100.0	379	19 W83962	Recombinant human
9	1263	100.0	379	19 W49073	Recombinant human
10	1263	100.0	396	18 W18574	Aggrecanase artifi
11	1263	100.0	396	18 W18575	Aggrecanase artifi
12	1263	100.0	400	21 Y15123	Porcine CTLA-4-Ig

13	1263	100.0	424	16 W14764	Human soluble kit
14	1263	100.0	424	16 W14765	Human soluble kit
15	1263	100.0	435	13 R26530	Sequence of one ch
16	1263	100.0	437	18 W10552	Alpha-1-acid glyco
17	1263	100.0	442	18 W10550	IgG1 polypeptide.
18	1263	100.0	445	20 Y24153	Bovine LOX-1 extra
19	1263	100.0	446	17 W05829	Humanised 1D10 ant
20	1263	100.0	447	20 Y31669	Human IgG1 chain C
21	1263	100.0	449	14 R43339	Completely humanis
22	1263	100.0	449	19 W49816	Amino acid sequenc
23	1263	100.0	452	20 Y30201	Heavy chain sequen
24	1263	100.0	459	14 R42066	Human anti-HBs hea
25	1263	100.0	460	21 Y69890	Human NR8a1pha/IgG
26	1263	100.0	461	14 R42162	Anti-HIV-1 recombi
27	1263	100.0	467	13 R22758	Reshaped C34 antib
28	1263	100.0	467	13 R22759	Reshaped C34 antib
29	1263	100.0	470	13 R22757	Reshaped C34 antib
30	1263	100.0	471	21 Y45030	HUMAN OCRI1-Fc fus
31	1263	100.0	475	13 R20057	Heavy chain of 3D6
32	1263	100.0	475	17 R93553	Monoclonal antibody
33	1263	100.0	475	18 W11641	Human anti-RSV mon
34	1263	100.0	475	18 W11639	Human anti-RSV mon
35	1263	100.0	476	14 R31023	Antibody D heavy c
36	1263	100.0	476	18 W01818	Primates anti-hu
37	1263	100.0	476	18 W01822	Primates anti-hu
38	1263	100.0	476	19 W63765	Macaque primatized
39	1263	100.0	476	19 W63761	Macaque primatized
40	1263	100.0	476	20 W88464	Monoclonal antibody
41	1263	100.0	477	20 W90207	hB7.2Fc soluble fu
42	1263	100.0	478	19 W63763	Macaque primatized
43	1263	100.0	480	20 W90206	hB7.1Fc soluble fu
44	1263	100.0	481	13 R24442	Sequence of antibo
45	1263	100.0	534	13 R26531	Sequence of CD4-Ig

#### ALIGNMENTS

RESULT 1

W26232 W26232 standard; Protein; 232 AA.

XX AC W26232;

XX AC W26232;

DT 16-MAR-1998 (first entry)

DE Human IgG1 hinge/Fc region.

XX KW Fusion protein; hydrophilic spacer; recombinant; expression system;

XX KW carboxypeptidase; IgG1; immunoglobulin; hinge region; Fc.

XX OS Homo sapiens.

XX PN WO9728272-A1.

XX PD 07-AUG-1997.

PF 31-JAN-1997; 97WO-US01470.

PR 31-JAN-1996; 96US-0595043.

XX (TECH-) TECHNOLOGENE INC.

XX PI Sgarlato GD;

XX DR WPI; 1997-402624/37.

XX DR N-PSDB; T80158.

XX PT Recombinant protein expression system for fusion protein production

XX PT - useful for high quantity production of authentic recombinant

XX proteins

XX Example 3; Page 133-134; 194pp; English.

XX A novel recombinant vector has been developed which comprises a  
 CC nucleotide sequence encoding a fusion protein. The fusion protein  
 CC comprises three domains joined together in order, from N-terminus to  
 CC C-terminus, of a first domain comprising a protein of interest, a second  
 CC domain comprising a hydrophilic spacer and an affinity domain, each  
 CC domain comprising amino acid residues. The present sequence represents  
 CC the hinge/Fc region of human IgG1, used in example 3 of the present  
 CC invention. The recombinant vector is used for the production of  
 CC authentic recombinant proteins of interest. The method of the invention  
 CC is useful for the expression of fusion proteins capable of isolation by  
 CC affinity chromatography in pro- or eukaryotic cells. This method allows  
 CC for the efficient cleavage and generation of authentic proteins of  
 CC interest that do not contain extraneous (i.e. non-naturally occurring)  
 CC amino acids.

XX Sequence 232 AA;

Query Match 100.0%; Score 1263; DB 18; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 1e-93;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPKSCDKTHTCPPAPPELLGGPSVFLPFPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
 DB 1 epksckdtkhtccpccpapeallggpsvflfpkpkdtlmisrtpevtcvvvdvshedpevkf 60  
 QY 61 NWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120  
 DB 61 nwydgvgevhnaaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiekt 120  
 QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 180  
 DB 121 iskakgqprepqvyltppsrdeltnqvsltclvkgfypsdiavewesngqpennykttp 180  
 QY 181 PVLSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 232  
 DB 181 pvlsgdgsfflyskltdvdkstrwqgnvfscsvmhcalhnhytqkslsispkg 232

RESULT 2  
 Y24154  
 ID Y24154 standard; Protein; 259 AA.  
 AC Y24154;  
 DT 10-SEP-1999 (first entry)  
 DE Protein from pCd5lneg1 comprising human IgG1 Fc region genomic DNA.  
 KW LDL; denatured; oxidized; arteriosclerosis; hyperlipidaemia;  
 KW low density lipoprotein; receptor; detection; immunoglobulin;  
 KW fusion protein.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN W09932520-A1.  
 PD 01-JUL-1999.  
 XX 18-DEC-1998; 98WO-05744.  
 XX 16-DEC-1998; 98JP-0358170.  
 PR 19-DEC-1997; 97JP-0364981.  
 PR 09-DEC-1998; 98JP-0349648.  
 XX (NISH) JAPAN TOBACCO INC.  
 XX Kakutani M, Masaki T, Sawamura T;  
 XX WPI; 1999-418906/35.  
 DR N-PSDB; X88533.

XX Fusion peptide for assay of oxidized LDL and for therapeutic use  
 PT Example 1; Page 92-96; 105pp; Japanese.  
 PS The present invention describes a fusion peptide which consists of the  
 CC extracellular domain of a mammalian oxidized LDL (low density  
 CC lipoprotein) receptor, fused to a partial heavy chain of a mammalian  
 CC immunoglobulin containing all or part of the constant region. Oxidized  
 CC LDL is a denatured form of LDL occurring in patients having  
 CC arteriosclerosis or hyperlipidaemia, and the fusion peptide can be  
 CC used for the assay of oxidized LDL in biological samples from such  
 CC patients, for the diagnosis of the disorders. It can also be used  
 CC therapeutically for the prevention and treatment of arteriosclerosis and  
 CC hyperlipidaemia. The present sequence represents the protein from the  
 CC vector DNA of pcd5lneg1 comprising human IgG1 Fc region genomic DNA.  
 XX Sequence 259 AA;  
 SQ Query Match 100.0%; Score 1263; DB 20; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-93;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPKSCDKTHTCPPAPPELLGGPSVFLPFPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
 DB 28 epksckdtkhtccpccpapeallggpsvflfpkpkdtlmisrtpevtcvvvdvshedpevkf 87  
 QY 61 NWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120  
 DB 88 nwydgvgevhnaaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiekt 147  
 QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTP 180  
 DB 148 iskakgqprepqvyltppsrdeltnqvsltclvkgfypsdiavewesngqpennykttp 207  
 QY 181 PVLSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 232  
 DB 208 pvlsgdgsfflyskltdvdkstrwqgnvfscsvmhcalhnhytqkslsispkg 259  
 RESULT 3  
 R91806  
 ID R91806 standard; Protein; 329 AA.  
 AC R91806;  
 XX 20-SEP-1996 (first entry)  
 DT Human immunoglobulin gamma heavy chain constant region sequences.  
 DE alkaline phosphatase; label; antibody; IgG; H-chain; C-region; CHI;  
 KW CH2; CH3; hinge; fusion protein; chimera; immunoassay.  
 XX Homo sapiens.  
 OS JP08070875-A.  
 PN 19-MAR-1996.  
 PD 05-SEP-1994; 94JP-0211035.  
 XX 05-SEP-1994; 94JP-0211035.  
 PR (TOYJ) TOSOH CORP.  
 PA WPI; 1996-203155/21.  
 XX N-PSDB; T27385.  
 DR Recombinant alkaline phosphatase (AP)-antibody fusion protein -  
 PT comprises AP fused downstream of antibody heavy or light chain,  
 PT useful as immunoassay reagent  
 XX

PS Example 1; Page 13-15; 44pp; Japanese.

XX The gene coding for human alkaline phosphatase is fused downstream

CC of a gene coding for either the variable and CH1 regions of an

CC antibody heavy chain or an antibody light chain. Coexpression of the

CC H- and L-chain sequences, one of which is fused to the AP gene,

CC results in production of AP-labelled antibodies suitable for use in

CC immunoassays. The present sequence is from a human IgG heavy chain

CC constant region.

XX Sequence 329 AA;

SY Query Match 100.0%; Score 1263; DB 17; Length 329;

Best Local Similarity 100.0%; Pred. No. 1.6e-93;

Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60

DB 98 epksckdtkhtcpcpapellgspvflfpkpkdtlmisrptevtcvvdvshedpevkf 157

QY 61 NWYVDGVEVHNKTPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKT 120

DB 158 nwyvdgvevhnaktpreeqynstyrsvltvlhqdwlngkeyckvsnkalpapiekt 217

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEESNGQPENNYKTTTP 180

DB 218 iskakgqprepqvtytlppsrdeltknqvsltcclvkgyfypsdiaveesngqpennyykttp 277

QY 181 PVLDSGSPFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSPGK 232

DB 278 pvltdsgsfflyskltvdksrwggnvfscsvmhealhnhytkslspgk 329

RESULT 4

ID R43685 standard; Protein; 351 AA.

AC R43685;

DT 25-MAY-1994 (first entry)

XX Human kappa immunoglobulin light chain constant domain.

DE Human; immunoglobulin; constant; region; humanised; P-selectin; light;

KW blocking; antibody; heavy; chain; variable; murine; thrombotic disease;

KW monoclonal; PBL.3; CDR; complementarity determining region; leukocyte;

KW expression vector; coexpression; pHCMV-1748RHA-gammalci-dhfr; epitope;

KW pHCMV-1748RHA-KR-neo; PBL.3/Humanised version A; vascular endothelium;

KW pHCMV-1747CH-gammaCi-neo; pHCMV-1747-CL-KR-neo; PBL.3 chimera;

KW acute lung injury; ischaemia reperfusion injury; inflammation.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Domain 22..119

FT Region /note= "CH1 domain"

FT Region 120..134

FT Domain /note= "Hinge region"

FT Domain 135..244

FT Domain /note= "CH2 domain"

FT Domain 245..352

FT Domain /note= "CH3 domain"

XX W09321956-A.

PN 11-NOV-1993.

XX 04-MAY-1993; 93WO-US04274.

XX 05-MAY-1992; 92US-0880196.

XX (CYTE-) CYTEL CORP.

XX Chestnut RW, Paulson JC, Polley MJ;

XX WPI: 1993-368423/46.

DR N-PSDB; Q51547.

XX Anti-P-selectin antibody for ischaemia acute lung injury treatment -

PT useful to treat inflammation and pathological conditions of

PT intercellular adhesion by competitive inhibition assays

XX Example 10; Fig 9; 82pp; English.

XX The sequences given in R43685-86 represent human immunoglobulin

CC constant regions which were used in the production of the humanised

CC P-selectin blocking antibody, along with the heavy and light chain

CC variable region coding sequences of the murine monoclonal antibody

CC PBL.3, given in R43687-88. The CDRs from PBL.3 heavy and light

CC chains were substituted for the CDRs of human heavy and light chains.

CC The humanised variable regions were inserted into expression vectors.

CC By coexpression of appropriate combinations of heavy and light

CC chains, several humanised antibodies can be expressed. Coexpression

CC of pHCMV-1748RHA-gammalci-dhfr and pHCMV-1748RHA-KR-neo gives rise

CC to the PBL.3/Humanised version A. Coexpression of pHCMV-1747CH-

CC gammaCi-neo and pHCMV-1747-CL-KR-neo gives rise to the PBL.3 chimera.

CC These humanised antibodies selectively bind epitopes on P-selectin and

CC block adhesion of leukocytes to the vascular endothelium. They may be

CC used to treat inflammatory and thrombotic diseases and other

CC pathological conditions involving P-selectin and antibodies to it, esp.

CC acute lung injury and ischaemia reperfusion injury.

XX Sequence 351 AA;

SY Query Match 100.0%; Score 1263; DB 14; Length 351;

Best Local Similarity 100.0%; Pred. No. 1.7e-93;

Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60

DB 120 epksckdtkhtcpcpapellgspvflfpkpkdtlmisrptevtcvvdvshedpevkf 179

QY 61 NWYVDGVEVHNKTPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKT 120

DB 180 nwyvdgvevhnaktpreeqynstyrsvltvlhqdwlngkeyckvsnkalpapiekt 239

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEESNGQPENNYKTTTP 180

DB 240 iskakgqprepqvtytlppsrdeltknqvsltcclvkgyfypsdiaveesngqpennyykttp 299

QY 181 PVLDSGSPFLYSLKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSPGK 232

DB 300 pvltdsgsfflyskltvdksrwggnvfscsvmhealhnhytkslspgk 351

RESULT 5

ID P91918 standard; protein; 371 AA.

XX AC P91918;

XX 14-MAY-1990 (first entry)

XX Sequence of the linked immunoglobulin gamma chain fragment.

XX Immunoglobulin gamma chain; IgG1 heavy chain constant region.

XX Key Location/Qualifiers

FT Misc-difference 42..43

FT /note= "Insert site"

FT Misc-difference 144..145

FT /note= "Insert site"

XX EP314317-A.

```
XX PD 03-MAY-1989.
XX PF
XX PR 03-OCT-1988; 88EP-0309194.
XX PR 28-SEP-1988; 88US-0250785, US-104329.
XX PA (GETH ) GENENTECH INC.
XX PI Capon DJ, Gregory TJ;
XX PS WPI; 1989-131855/18.
XX DR N-PSDB; N90779.
XX CC Compsns. contg. ahdesion variants
XX PT - useful in therapy and diagnostics, eg CD4 variants
XX PT which are therapeutically useful for treating human
XX PT immuno-deficiency virus
XX PS Fig 4a-4b; : 36pp; English.
XX CC It may be fused to the first 180 N-terminal
XX CC residues of CD4 at the C-terminus. The fusion protein may be used for
XX CC antiviral of immunomodulatory therapy particularly in treatment of HIV
XX CC infection.
XX SQ Sequence 371 AA;

Query Match 100.0%; Score 1263; DB 10; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
DB 140 epkscdkhtcpcpapellggpsvflfppkpkdtlmisrtpevtcvvvdvshedpevkf 199
QY 61 NWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120
DB 200 nwydgvghvhnaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiekt 259
QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTP 180
DB 260 iskakgqprepqvylppsrdeltknqvscltclvkgyfypsdiavwesngqpennykttp 319
QY 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGK 232
DB 320 pvlsgdsfflyskltvdksrwqgnvfscvmhealhnhytqkslsispgk 371

RESULT 6
ID P93558 standard; protein; 371 AA.
XX AC P93558;
XX DT 06-JUN-1990 (first entry)
XX DE Linkered human IgG1 (gamma 1) chain fragment.
XX KW Human IgG1; gamma 1; immunoglobulin; CD4; fusion protein.
XX OS Homo sapiens.
XX PN W08902922-A.
XX PD 06-APR-1989.
XX PF 03-OCT-1988; 88WO-US03414.
XX PR 02-OCT-1987; 87US-0104329.
XX PA (GETH ) GENENTECH.
```

```
XX PI Capon DJ, Gregory TJ;
XX DR WPI; 1989-114397/15.
XX DR N-PSDB; N90736.
XX PT New nucleic acid sequences encoding adhesion, esp. CD4, variants -
XX PT partic. with trans-membrane domain inactivated or fused to other
XX PT peptide, useful esp. for treating HIV infections
XX PS Fig 4A-4B2; pp. 10/13-12/13; 78pp; English.
XX CC It is employed in the prepn. of CD4 fusions. CD4 fusion proteins can
XX CC have antiviral and immunomodulatory activity and are esp. useful for
XX CC treating HIV infections, regardless of genetic variation within the
XX CC virus. They and antibodies raised against them can also be used
XX CC diagnostically for assaying adhesions and their ligands.
XX SQ Sequence 371 AA;

Query Match 100.0%; Score 1263; DB 10; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.8e-93;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60
DB 140 epkscdkhtcpcpapellggpsvflfppkpkdtlmisrtpevtcvvvdvshedpevkf 199
QY 61 NWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120
DB 200 nwydgvghvhnaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiekt 259
QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTP 180
DB 260 iskakgqprepqvylppsrdeltknqvscltclvkgyfypsdiavwesngqpennykttp 319
QY 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGK 232
DB 320 pvlsgdsfflyskltvdksrwqgnvfscvmhealhnhytqkslsispgk 371

RESULT 7
ID W60037 standard; Protein; 376 AA.
XX AC W60037;
XX DT 11-SEP-1998 (first entry)
XX DE Antigenic peptide hFas (nd29) containing FC region.
XX KW Fas ligand; Fas antagonist; apoptosis related disease; liver disease;
XX KW heart failure; kidney failure; graft-versus-host disease; antibody;
XX KW myocardial infarction; ischemic restenosis; endotoxic shock.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Peptide 1..16
XX FT Protein /note= "hFas antigen signal peptide"
XX FT /note= "hFas (nd29) protein"
XX PN W09818487-A1.
XX PD 07-MAY-1998.
XX PF 31-OCT-1997; 97WO-JP03978.
XX PR 26-SEP-1997; 97JP-0262521.
XX PR 31-OCT-1996; 96JP-0290459.
XX PR 27-DEC-1996; 96JP-0351718.
```



XX (MOCH ) MOCHIDA PHARM CO LTD.  
 PA (OSAB-) OSAKA BIOSCIENCE INST.  
 XX Nagata S, Suda T, Yatomi T;  
 XX WPI; 1998-271925/24.  
 DR N-PSDB; V34430.  
 XX  
 XX Use of Fas antagonist for treatment and prevention of  
 PT apoptosis-related diseases - such as heart or kidney failure,  
 PT graft-versus-host disease or liver disease  
 XX  
 PS Examples; Fig 5-9; 86pp; Japanese.  
 XX  
 CC This represents the antigenic peptide hFas (nd29) containing the Fc  
 CC region. The invention provides the use of Fas antagonist as an agent for  
 CC the treatment and prevention of apoptosis-related diseases. The Fas  
 CC antagonist can be a partial Fas antigen peptide containing the  
 CC extracellular part of the protein, but lacking the signal sequence, an  
 CC anti-Fas antibody, or an anti-Fas ligand antibody, where the antibody is  
 CC preferably a humanised antibody. The Fas antagonist is used in the  
 CC treatment and prevention of diseases such as myocardial infarction, heart  
 CC failure, ischemic heart disease, acute kidney failure, graft-versus-host  
 CC disease, ischemic restenosis of the heart, liver or kidney, and  
 CC endotoxic shock, and also as an organ preservative in transplantation.  
 CC The agent is of low toxicity but effectively inhibits the Fas/Fas ligand  
 CC system.  
 XX  
 SQ Sequence 376 AA;

Query Match 100.0%; Score 1263; DB 19; Length 376;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-93;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60  
 DB 145 epkscdkthtcpcpapellgpgsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 204  
 QY 61 NNYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120  
 DB 205 nnyvdgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiekt 264  
 QY 121 ISKAKGQPEPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVENSGQPNENYKTTTP 180  
 DB 265 iskakgqpepvytlppsrdeltnqvsltcclvkgfypsdiavensngqpennykttcp 324  
 QY 181 PVLSDSGSFLLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 232  
 DB 325 pvlsgdgsfllyskltvdksrwqgnvfscsvmhealhnhytqkslspsgk 376

RESULT 8  
 ID W83962 standard; Protein; 379 AA.  
 XX AC W83962;  
 XX  
 DT 15-FEB-1999 (first entry)  
 XX Recombinant human metFc-OB protein.  
 DE  
 KW Recombinant; metFc-OB protein; Fc region; immunoglobulin; Ig; OB;  
 KW obesity; human; adiposity; blood lipid; diabetes type II; insulin;  
 KW hypoglycaemic; antihypertensive; diuretic; appetite suppressant;  
 KW suspension.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 5  
 FT /note= "can be optionally replaced with Ala"

FT Misc-difference 20  
 FT /note= "can be optionally replaced with Glu"  
 FT Misc-difference 103  
 FT /note= "can be optionally replaced with Ala"  
 FT Misc-difference 105  
 FT /note= "can be optionally replaced with Ala"  
 FT Misc-difference 107  
 FT /note= "can be optionally replaced with Ala"  
 XX  
 PN W09846257-Al.  
 XX  
 XX 22-OCT-1998.  
 PD  
 XX  
 PF 16-APR-1998; 98WO-US07828.  
 XX  
 PR 14-APR-1998; 98US-0059467.  
 PR 17-APR-1997; 97US-0843971.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Brens DN, French DL, Speed MA;  
 XX  
 DR WPI; 1998-594525/50.  
 DR N-PSDB; V69685.  
 XX  
 CC Concentrated suspension of fusion of obesity protein with Fc  
 CC immunoglobulin fragment - stable at physiological pH, used for e.g.  
 CC reduction of weight and blood lipid levels, and for treatment of  
 CC type II diabetes  
 PT  
 PT  
 XX  
 PS Claim 2; Fig 5A-C; 47pp; English.  
 XX  
 CC This represents a recombinant metFc-OB protein which consists of an Fc  
 CC region of human immunoglobulin (Ig) fused to a human OB (obesity)  
 CC protein. The invention provides a human OB protein suspension that  
 CC contains at least 0.5 mg/ml of the human OB protein derivatised by  
 CC attachment of the Fc region of an Ig to the N-terminus of OB, and has a  
 CC pH 6-8. The suspensions are used to reduce weight, adiposity and blood  
 CC lipid levels, to treat or prevent diabetes type II, and to increase lean  
 CC mass and insulin sensitivity. They may be used in conjunction with  
 CC insulin, hypoglycaemics, antihypertensives, diuretics, appetite  
 CC suppressants etc. These suspensions are stable and active at  
 CC physiological pH and are ready-for-use formulations that do not require  
 CC freezing or freeze drying. As they are very concentrated, only small  
 CC volumes are required and they provide a sustained-release effect, with  
 CC increased potency and reduced frequency of injection.  
 XX  
 SQ Sequence 379 AA;

Query Match 100.0%; Score 1263; DB 19; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-93;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60  
 DB 2 epkscdkthtcpcpapellgpgsvflfppkpkdtlmisrtpevtcvvvdshedpevkf 61  
 QY 61 NNYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120  
 DB 62 nnyvdgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeykckvsnkalpapiekt 121  
 QY 121 ISKAKGQPEPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVENSGQPNENYKTTTP 180  
 DB 122 iskakgqpepvytlppsrdeltnqvsltcclvkgfypsdiavensngqpennykttcp 181  
 QY 181 PVLSDSGSFLLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 232  
 DB 182 pvlsgdgsfllyskltvdksrwqgnvfscsvmhealhnhytqkslspsgk 233

RESULT 9  
 W49073

ID W49073 standard; Protein; 379 AA.  
XX W49073;  
XX  
XX  
XX 18-NOV-1998 (first entry)  
XX  
XX Recombinant human MetFc-OB protein.  
DE  
XX  
XX Recombinant human MetFc-OB protein; chimeric; immunoglobulin; diabetes;  
KW high blood lipid level; arterial sclerosis; stroke; Fc-OB fusion protein.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX  
XX Key Location/Qualifiers  
FH Protein 2...379 "Recombinant human Fc-OB protein"  
FT /note= 234...379  
FT Region /note= "Human OB protein"  
XX  
XX WO9828427-A1.  
XX  
XX 02-JUL-1998.  
PD  
XX 11-DEC-1997; 97WO-US23183.  
PF  
XX 20-DEC-1996; 96US-0770973.  
PR  
XX (AMGE-) AMGEN INC.  
PA  
XX Hecht RI, Mann MB;  
PI  
XX WPI; 1998-377658/32.  
DR N-PSDB; V32900.  
XX  
XX New fusion proteins of OB and Fc - used for treating e.g. excess  
PT weight, diabetes, arterial sclerosis, arterial plaque, high blood  
PT lipid level, gall stones or stroke  
XX  
XX Claim 2; Fig 3A-3C; 107pp; English.  
XX  
XX The present sequence represents a recombinant human MetFc-OB fusion  
CC protein. The invention provides Fc-OB fusion proteins whereby the  
CC Fc region of an immunoglobulin or its analogue is linked, either directly  
CC or indirectly using a linker, to the N-terminus of an OB protein or its  
CC analogue. The Fc-OB fusion proteins are claimed to demonstrate increased  
CC stability and clearance rate and decreased degradation as compared to OB  
CC protein or a fusion of Fc to the C-terminus of the OB protein. These  
CC Fc-OB fusion proteins are also claimed to be useful for treating excess  
CC weight in an individual or animal or for treating co-morbidities  
CC associated with excess fat such as diabetes, high blood lipid level,  
CC arterial sclerosis and stroke.  
XX  
XX Sequence 379 AA;  
SQ

Query Match 100.0%; Score 1263; DB 19; Length 379;  
Best Local Similarity 100.0%; Pred. No. 1.8e-93;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EPKSCDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
Db 2 epkscdkthtccpcapellggpsvflfpkpkdtlmisrtpevtcvvvdshedpevkf 61  
QY 61 NWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKT 120  
Db 62 nwyvdgvevhnatkpreedynstyrsvsvltvlhqdwlngkeyckvsnkalpapiekt 121  
QY 121 ISKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGFPVPSDIAVEWESNGQPENNYKTP 180  
Db 122 iskakgprepqvylppsrdeltknqvsltcclvkgfypsdiavewesngqpennnyktp 181  
QY 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFSVMHEALHNHYTQKSLSLSPCK 232

Db 182 pvlsgdsfflyskltvdksrwqgnvfscvmhealhnhytqkslspspk 233  
RESULT 10  
W18574  
ID W18574 standard; Protein; 396 AA.  
XX  
XX W18574;  
AC  
XX 17-SEP-1997 (first entry)  
DT  
XX  
XX Aggrecanase artificial recombinant substrate ragg-1.  
DE  
XX Artificial recombinant substrate; ragg1; aggrecanase; aggrecan;  
KW osteoarthritis; diagnosis.  
KW  
XX Chimaeric Homo sapiens;  
OS Chimaeric synthetic.  
OS  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..24  
FT /label= Sig-peptide  
FT /note= "CD5 signal sequence"  
FT Peptide 25..32  
FT /label= FLAG  
FT Domain 33..160  
FT /label= Aggrecan  
FT /note= "human aggrecan interglobular domain"  
FT Peptide 161..164  
FT /label= Spacer  
FT Region 165..179  
FT /label= Hinge  
FT /note= "human IgG1 hinge region"  
FT Region 180..289  
FT /label= CH2  
FT /note= "human IgG1 CH2 region"  
FT Region 290..396  
FT /label= CH3  
FT /note= "human IgG1 CH3 region"  
XX  
XX EP785274-A1.  
PN  
XX 23-JUL-1997.  
PD  
XX 27-DEC-1996; 96EP-0120949.  
PF  
XX 18-JAN-1996; 96EP-0100682.  
PR  
XX (FARH ) HOECHST AG.  
PA  
XX Bartnik E, Buettner F, Caterson B, Eidenmueller B;  
PI Hughes C;  
PI  
XX  
XX WPI; 1997-365948/34.  
DR N-PSDB; T69892.  
DR  
XX Recombinant substrate for aggrecanase in vitro testing - and  
PT encoding DNA, useful for studying aggrecanase activity e.g. by  
PT detection of cleavage products for monitoring onset or progression  
PT of osteoarthritis  
XX  
XX Claim 3; Page 15-16; 28pp; English.  
PS  
XX An artificial recombinant substrate, ragg-1 (W18574), for  
CC aggrecanase comprises the CD5 signal sequence, a FLAG epitope for  
CC M1 monoclonal antibody detection, the interglobular domain of human  
CC aggrecan, and human IgG1 hinge, CH2 and CH3 regions. It is the  
CC expression product of a DNA molecule (T69892) that can be  
CC incorporated into a vector for use in ragg-1 prodn. in host cells.  
CC ragg-1 can be used in cell culture systems to study the activity of  
CC aggrecanase, to detect new enzymatic cleavage sites, for the  
CC affinity purification of aggrecanase, to isolate aggrecanase

CC cDNA by functional cloning, to screen for aggrecanase inhibitors,  
CC in methods for monitoring the onset or progression of  
CC osteoarthritis, and in diagnostic aids. Another rAGG-1 (W18575)  
CC has alanine at amino acid position 34.  
XX Sequence. 396 AA;  
SQ

Query Match 100.0%; Score 1263; DB 18; Length 396;  
Best Local Similarity 100.0%; Pred. No. 1.9e-93;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
DB 165 epksckdtkthtccpapepallggpsvflfppkpkdtlmisrtpvctcvvvdshedpevkf 224

QY 61 NWYVDGVEVHNKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120  
DB 225 nwydgvvevhnaktprreeqynstyrsvsvltvlhqdwlngkeykckvsnkalpapiekt 284

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 180  
DB 285 iskagqprepqvytlppsrdeltnkvsltcvkgyfypsdiavewesngqpennyykttp 344

QY 181 PVLDSGSPFLYSLKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGK 232  
DB 345 pvlldsgsflyslkltvdksrwqgnvfscvmhealhnhytqkslsispgk 396

RESULT 11  
W18575  
ID W18575 standard; Protein; 396 AA.  
XX  
AC W18575;  
XX  
DT 17-SEP-1997 (first entry)  
XX  
DE Aggrecanase artificial recombinant substrate rAGG-1.  
XX  
KW Artificial recombinant substrate; rAGG1; aggrecanase; aggrecan;  
KW Osteoarthritis; diagnosis.  
XX  
OS Chimaeric Homo sapiens;  
OS Chimaeric synthetic.  
FH  
FH Key Location/Qualifiers  
FT Peptide 1..24  
FT /label= sig\_peptide  
FT /note= "CD5 signal sequence"  
FT Peptide 25..32  
FT /label= FLAG  
FT Domain 33..160  
FT /label= Aggrecan  
FT /note= "human aggrecan interglobular domain"  
FT Peptide 161..164  
FT /label= Spacer  
FT Region 165..179  
FT /label= Hinge  
FT /note= "human IgG1 hinge region"  
FT Region 180..289  
FT /label= CH2  
FT /note= "human IgG1 CH2 region"  
FT Region 290..396  
FT /label= CH3  
FT /note= "human IgG1 CH3 region"

EP785274-A1.  
XX  
PN  
XX  
PD  
XX  
PF 27-DEC-1996; 96EP-0120949.  
XX  
PR 18-JAN-1996; 96EP-0100682.

(FARH ) HOECHST AG.  
XX  
PA  
XX Bartnik E, Buettner F, Caterson B, Eidenmueller B;  
PI Hughes C;  
XX  
DR WPI; 1997-365948/34.  
DR N-PSDB; T69893.  
XX  
XX Recombinant substrate for aggrecanase in vitro testing - and  
PT encoding DNA, useful for studying aggrecanase activity e.g. by  
PT detection of cleavage products for monitoring onset or progression  
PT of osteoarthritis  
XX  
PS Claim 3; Refer to page 15-16; 28pp; English.  
XX  
CC An artificial recombinant substrate, rAGG-1 (W18575), for  
CC aggrecanase comprises the CD5 signal sequence, a FLAG epitope for  
CC M1 monoclonal antibody detection, the interglobular domain of human  
CC aggrecan, and human IgG1 hinge, CH2 and CH3 regions. It is the  
CC expression product of a DNA molecule (T69893) that can be  
CC incorporated into a vector for use in rAGG-1 prodn. in host cells.  
CC rAGG-1 can be used in cell culture systems to study the activity of  
CC aggrecanase, to detect new enzymatic cleavage sites, for the  
CC affinity purification of aggrecanase, to isolate aggrecanase  
CC cDNA by functional cloning, to screen for aggrecanase inhibitors,  
CC in methods for monitoring the onset or progression of  
CC osteoarthritis, and in diagnostic aids. Another rAGG-1 (W18574)  
CC has glycine at amino acid position 34.  
XX  
SQ Sequence 396 AA;

Query Match 100.0%; Score 1263; DB 18; Length 396;  
Best Local Similarity 100.0%; Pred. No. 1.9e-93;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCPPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
DB 165 epksckdtkthtccpapepallggpsvflfppkpkdtlmisrtpvctcvvvdshedpevkf 224

QY 61 NWYVDGVEVHNKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120  
DB 225 nwydgvvevhnaktprreeqynstyrsvsvltvlhqdwlngkeykckvsnkalpapiekt 284

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 180  
DB 285 iskagqprepqvytlppsrdeltnkvsltcvkgyfypsdiavewesngqpennyykttp 344

QY 181 PVLDSGSPFLYSLKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGK 232  
DB 345 pvlldsgsflyslkltvdksrwqgnvfscvmhealhnhytqkslsispgk 396

RESULT 12  
Y15123  
ID Y15123 standard; Protein; 400 AA.  
XX  
AC Y15123;  
XX  
DT 07-FEB-2000 (first entry)  
XX  
DE Porcine CTLA-4-Ig construct.  
XX  
KW Porcine CTLA-4; soluble protein; xenograft; organ transplant; B7; CD28;  
KW xenograft-specific immunosuppression; recipient T-cell; anergy;  
KW co-stimulatory signal 2; homology; human CTLA-4; bovine CTLA-4.  
OS Sus scrofa.  
OS Homo sapiens.  
FH  
FH Key Location/Qualifiers  
FT Region 162..168

FT FT /label= Flexible\_linker  
 FT FT /note= "Denotes the junction between pCTLA-4"  
 FT Domain 169..362  
 FT /label= IgG1\_domain

XX WO9957266-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-GB01350.

XX 30-APR-1998; 98GB-0009280.

XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.

XX Lechler IR, Dorling A;

XX WPI; 2000-038815/03.

XX Inhibiting T-cell mediated rejection of xenotransplanted organs

XX Claim 1; Fig 4; 43pp; English.

CC The present sequence is porcine CTLA-4-Ig construct for xenograft  
 CC -specific immunosuppression. In a pig-to-human transplantation, the  
 CC soluble protein could comprise the extracellular domain of porcine CTLA-4  
 CC fused to a human gamma 1 chain of IgG1. This construct was subcloned  
 CC into the expression vector pHOOK-3TM and used to transfect DAP.3 or  
 CC CHO-K1 cells. pCTLA-4-Ig preferentially binds to porcine B7 and blocks  
 CC its interaction with CD28 on recipient T-cells. This is useful as a  
 CC species-specific reagent to inhibit human T-cell proliferative responses  
 CC to a variety of stimulators.

XX Sequence 400 AA;

Query Match 100.0%; Score 1263; DB 21; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 2e-93;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60

Db 169 epksdcthtctcpapellggpsvflfppkpkdtlmisrtpetvctvvdvshedpevkf 228

QY 61 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120

Db 229 nwyvdgvevhnakckpreedynstyrvvsvltvlnhqdwnlqngkeyckvsnkalpapiekt 288

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 180

Db 289 iskakgqprepqvylppsrdelctknqvsclclvkgyfypsdiawesngqpennnyktp 348

QY 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKSLSLSPGK 232

Db 349 pvlsgdsfflyskltvdkswrggnvfscsvmhealhnhytqkslsispkg 400

RESULT 13

W14764

ID W14764 standard; Protein; 424 AA.

XX AC

XX W14764;

XX 11-JUN-1997 (first entry)

XX Human soluble kit ligand-IgG fusion protein.

DE Kit ligand; c-kit proto-oncogene; cytokine; growth factor;  
 KW haematopoietic cell; cell proliferation; stem cell; anaemia;  
 KW thrombocytopaenia; therapy; IgG1.

XX Homo sapiens.

DE Human soluble kit ligand-IgG fusion protein (corrected).

XX Key Location/Qualifiers  
 FH Peptide 1..25  
 FT /label= Sig\_peptide  
 FT /note= "KL signal peptide"  
 FT Protein 26..424  
 FT /label= Mat\_protein  
 FT /note= "human KL-IgG fusion"

XX WO9526199-A1.

XX 05-OCT-1995.

XX 28-MAR-1995; 95WO-US03866.

XX 28-MAR-1994; 94US-0220379.

XX (CYTO-) CYTOMED INC.

XX Lobell RB, Nocka KH;

XX WPI; 1995-351198/45.

XX N-PSDB; T63109.

XX Covalent dimers of kit ligand or FLT-3/FLK-2 ligand - exhibit  
 XX increased activity in promoting cell proliferation

XX Claim 10; Page 43-44; 88pp; English.

CC A fusion protein (W14764) between human soluble kit ligand (KL)  
 CC (see also W14761) and a human IgG1 heavy chain can be transiently  
 CC expressed in COS cells transfected with a human KL-Ig cDNA  
 CC construct (T63109) in vector CDM8; a corrected KL-Ig construct  
 CC (W14765) has also been prepd. KL-Ig can also be produced as a  
 CC dimer stabilised by intermolecular disulphide bonds or a peptide  
 CC linker. The stabilised KL-Ig dimers have a more favorable cell  
 CC proliferation: mast cell activation ratio than native KL and can  
 CC stimulate haematopoietic recovery or stem cell/progenitor cell  
 CC mobilisation with less toxicity.

XX Sequence 424 AA;

Query Match 100.0%; Score 1263; DB 16; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-93;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKF 60

Db 193 epksdcthtctcpapellggpsvflfppkpkdtlmisrtpetvctvvdvshedpevkf 252

QY 61 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120

Db 253 nwyvdgvevhnakckpreedynstyrvvsvltvlnhqdwnlqngkeyckvsnkalpapiekt 312

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 180

Db 313 iskakgqprepqvylppsrdelctknqvsclclvkgyfypsdiawesngqpennnyktp 372

QY 181 PVLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKSLSLSPGK 232

Db 373 pvlsgdsfflyskltvdkswrggnvfscsvmhealhnhytqkslsispkg 424

RESULT 14

W14765

ID W14765 standard; Protein; 424 AA.

XX AC

XX W14765;

XX 11-JUN-1997 (first entry)

XX Kit ligand; c-kit proto-oncogene; cytokine; growth factor;  
 KW haematopoietic cell; cell proliferation; stem cell; anaemia;  
 KW thrombocytopaenia; therapy; IgG1.  
 XX Homo sapiens.

Key Location/Qualifiers  
 Peptide 1..25  
 /label= Sig\_peptide  
 /note= "KL signal peptide"  
 Protein 26..424  
 /label= Mat\_protein  
 /note= "human KL-IgG fusion"

W09526199-A1.  
 05-OCT-1995.  
 28-MAR-1995; 95WO-US03866.  
 28-MAR-1994; 94US-0220379.

(CYTO-) CYTOMED INC.  
 Lobell RB, Nocka KH;  
 WPI: 1995-351198/45.  
 N-PSDB; T63110.

Covalent dimers of kit ligand or FLT-3/FLK-2 ligand - exhibit  
 increased activity in promoting cell proliferation  
 Claim 10; Page 46-48; 88pp; English.

A fusion protein (W14765) between human soluble kit ligand (KL)  
 (see also W14761) and a human IgG1 heavy chain can be transiently  
 expressed in COS cells transfected with a human KL-Ig CDNA  
 construct (T63110) in vector CDM8. KL-Ig can also be produced as a  
 dimer stabilised by intermolecular disulphide bonds or a peptide  
 linker. The stabilised KL-Ig dimers have a more favorable cell  
 proliferation: mast cell activation ratio than native KL and can  
 stimulate haematopoietic recovery or stem cell/progenitor cell  
 mobilisation with less toxicity.

Sequence 424 AA;

Query Match 100.0%; Score 1263; DB 16; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-93;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCPAPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
 DB 193 epkscdkthtccpcpapelggpsvflfpppkpdtlmisrtpevtcvvvdshedpevkf 252

QY 61 NWYDGVVEVHNKTKPREQYNSTYRVSVLTVLDHODWLNKGEYCKVSNKALPAPIEKT 120  
 DB 253 nwydgvvevhnaktkpreeqynstyrsvsvltvlhdwlngkeyckvsnkalpapiekt 312

QY 121 ISKAKGQPREPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 180  
 DB 313 iskakgqprepvtytlppsrdeltnqvsltcclvkgyfypsdiavevesngqpennykttt 372

QY 181 PVLDSGSEFLYSKLTVDKSRQQGNVFCSCVMHEALHNHYTQKSLSLSPGK 232  
 DB 373 pvltdsgsflyskltvdksrqgnvfscsvmhealhnhytqkslspsgk 424

RESULT 15  
 R26530  
 ID R26530 standard; Protein; 435 AA.  
 XX

R26530;  
 XX 28-JAN-1993 (first entry)  
 DE Sequence of one chain of a CD4-gamma 1 chimeric heavy chain  
 DE homodimer.  
 XX CD4-gamma 1 chimeric heavy chain homodimer; expression vector; HIV;  
 KW therapy; diagnostic agent; inhibition.  
 XX Synthetic.  
 Key Location/Qualifiers  
 Region 1..204  
 /label= CD4  
 /note= "1..25 = preregion"  
 Region 205..219  
 /label= hinge  
 Region 220..329  
 /label= CH2  
 Region 330..436  
 /label= CH3

W09213559-A.  
 20-AUG-1992.

10-FEB-1992; 92WO-US01152.  
 08-FEB-1991; 91US-0654205.  
 (PROG-) PROGENICS PHARM INC.

Beaudry GA, Maddon PJ;  
 WPI: 1992-299758/36.  
 N-PSDB; Q27830.

CD4-gamma 1 chimeric heavy chain homo-dimer and its expression  
 vector - for preventing and treating HIV infection useful as a  
 diagnostic agent

Example; Fig 3; 88pp; English.

Human CD4 cDNA was excised from pSP6T4 and cloned into M13mp18. The  
 2 kb PstI/PstI fragment from pBR lambda 1 contg. the human lambda 1  
 heavy chain gene (contg. the hinge, CH2 and CH3 exons) was isolated  
 and cloned into the BAP-treated M13mp18/CD4 vector. To obtain a  
 CD4-lambda 1 chimeric heavy chain gene, oligonucleotide-mediated  
 site-directed mutagenesis was performed to juxtapose the CD4 and  
 lambda 1 heavy chain DNA sequences, ligating the CD4 sequence in  
 frame to the hinge exon. The DNA was then cloned into pCDNA-1 to  
 produce CD4-IgG1-pcDNA1 (ATCC 40951).

Sequence 435 AA;

Query Match 100.0%; Score 1263; DB 13; Length 435;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-93;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCPAPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
 DB 204 epkscdkthtccpcpapelggpsvflfpppkpdtlmisrtpevtcvvvdshedpevkf 263

QY 61 NWYDGVVEVHNKTKPREQYNSTYRVSVLTVLDHODWLNKGEYCKVSNKALPAPIEKT 120  
 DB 264 nwydgvvevhnaktkpreeqynstyrsvsvltvlhdwlngkeyckvsnkalpapiekt 323

QY 121 ISKAKGQPREPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTP 180  
 DB 324 iskakgqprepvtytlppsrdeltnqvsltcclvkgyfypsdiavevesngqpennykttt 383



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:20:02 ; Search time 135.68 seconds  
(without alignments)  
346.406 Million cell updates/sec

Title: US-09-389-782a-2  
Perfect score: 2179  
Sequence: 1 MNKWLCCALLVLLDIETWT.....OKLFLEMIGNQVSKISCL 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTEMBL\_15.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2179	100.0	401	11	O08712
2	2079	95.4	401	11	O08727
3	1892	86.8	401	4	O00300
4	1789	82.1	372	4	Q9UHP4
5	424.5	19.5	300	4	O95407
6	390.5	17.9	302	13	Q9PUS0
7	314	14.4	459	11	Q62327
8	302.5	13.9	482	11	O08734
9	301.5	13.8	655	4	O75509
10	298	13.7	439	4	Q16042
11	268.5	12.3	625	11	O35305
12	266.5	12.2	616	4	Q9Y606
13	226	10.4	355	12	O85308
14	224	10.3	349	12	O57098
15	222.5	10.2	349	12	O57305
16	221	10.1	349	12	O57284
17	220	10.1	349	12	O57097
18	219	10.1	349	12	O57110
19	218	10.0	349	12	O57111

20	218	10.0	349	12	Q89118
21	217.5	10.0	349	12	O57099
22	216	9.9	349	12	O89098
23	215.5	9.9	348	12	O57112
24	215.5	9.9	348	12	O85407
25	215.5	9.9	349	12	O57291
26	215.5	9.9	349	12	O57100
27	215.5	9.9	349	12	O57101
28	215.5	9.9	349	12	O57102
29	214.5	9.8	360	12	O57118
30	214	9.8	350	12	O57116
31	212.5	9.8	347	12	O57119
32	212	9.7	348	12	O57277
33	212	9.7	348	12	O57103
34	212	9.7	348	12	O57108
35	211.5	9.7	326	12	O57120
36	210.5	9.6	347	12	O57115
37	209.5	9.6	350	12	O57123
38	209	9.6	349	12	O57109
39	207.5	9.5	351	12	O57117
40	207.5	9.5	351	12	O73559
41	205.5	9.4	326	12	O57122
42	204.5	9.4	351	12	O57121
43	204	9.4	283	4	Q92956
44	204	9.4	283	4	Q9UM65
45	204	9.4	320	12	O57091

#### ALIGNMENTS

RESULT	1
O08712	
ID	O08712 PRELIMINARY; PRT; 401 AA.
AC	O08712; O70202;
DT	01-JUL-1997 (Tremblrel. 04, Created)
DT	01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT	01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE	OSTEOCALCIN PRECURSOR (OSTEOCALCIN INHIBITORY FACTOR)
DE	(OCIF).
GN	TNFRSF11B OR OPF.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BALE/C; TISSUE=KIDNEY;
RX	MEDLINE=97262071; PubMed=9108485;
RA	Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA	Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,
RA	Derose M., Elliott R., Colombero A., Tan H.-L., Trall G., Sullivan J.,
RA	Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA	Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA	Suggs S., Boyle W.J.;
RT	"Osteoprotegerin: a novel secreted protein involved in the regulation
RT	of bone density.";
RL	Cell 89:309-319(1997).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=129/OLA, AND NIH SWISS;
RX	MEDLINE=98382527; PubMed=9714833;
RA	Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,
RA	Higashio K.;
RT	"Structure of the mouse osteocalcogenesis inhibitory factor (OCIF)
RT	gene and its expression in embryogenesis.";
RL	Gene 215:339-343(1998).
CC	- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC	OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC	SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC	OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC	STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC	- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,  
 CC BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND  
 CC PLACENTA. NOT DETECTED IN SPLEEN.  
 CC -1- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT  
 CC DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY  
 CC 15 TO DAY 17.  
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; U94331; AAB53708.1; -.  
 DR EMBL; U94331; AAB53708.1; -.  
 DR EMBL; AB013898; BAA28269.1; -.  
 DR EMBL; AB013903; BAA33388.1; -.  
 DR EMBL; AB013899; BAA33388.1; JOINED.  
 DR EMBL; AB013900; BAA33388.1; JOINED.  
 DR EMBL; AB013901; BAA33388.1; JOINED.  
 DR EMBL; AB013902; BAA33388.1; JOINED.  
 DR HSSP; P25942; ICDF.  
 DR MGD; MGI:109587; Opg.  
 DR INTERPRO; IPR000488; -.  
 DR INTERPRO; IPR001368; -.  
 DR PFAM; PF00020; TNFR\_C6; 3.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; PS0017; DEATH DOMAIN; 1.  
 DR PROSITE; PS0050; TNFR\_NGFR\_2; 2.  
 DR PRODOM; PD000771; -. 1.  
 KW Glycoprotein; Repeat; Cytokine; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 401 OSTEOPROTEGERIN.  
 FT DOMAIN 23 201 4 X TNFR-CYS.  
 FT REPEAT 23 63 TNFR-CYS 1.  
 FT REPEAT 64 106 TNFR-CYS 2.  
 FT REPEAT 107 143 TNFR-CYS 3.  
 FT REPEAT 144 201 TNFR-CYS 4.  
 FT DOMAIN 306 365 DEATH DOMAIN.  
 FT DISULFID 41 54 BY SIMILARITY.  
 FT DISULFID 44 62 BY SIMILARITY.  
 FT DISULFID 65 80 BY SIMILARITY.  
 FT DISULFID 83 97 BY SIMILARITY.  
 FT DISULFID 87 105 BY SIMILARITY.  
 FT DISULFID 118 142 BY SIMILARITY.  
 FT DISULFID 145 160 BY SIMILARITY.  
 FT CARBOHYD 98 98 POTENTIAL.  
 FT CARBOHYD 165 165 POTENTIAL.  
 FT CARBOHYD 178 178 POTENTIAL.  
 FT CARBOHYD 289 289 POTENTIAL.  
 FT VARIANT 138 138 SWISS.  
 FT VARIANT 161 161 I -> R (IN STRAINS 129/OLA AND NIH  
 FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH  
 FT VARIANT 165 165 S -> A (IN STRAINS 129/OLA AND NIH  
 FT VARIANT 288 288 SWISS).  
 FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH  
 FT VARIANT 296 296 SWISS).  
 FT SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;  
 Query Match Similarity 100.0%; Score 2179; DB 11; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 6.7e-164;  
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MNKWLCCALLVLLDIETWTOTETPPKYLHYDPETGHOLLCDKCAPGYLKHCHTVRKRT 60  
 Db 1 MNKWLCCALLVLLDIETWTOTETPPKYLHYDPETGHOLLCDKCAPGYLKHCHTVRKRT 60  
 Qy 61 LCVPCPDHSYDTSNHTSDCYVCKELQSVKQECNRTNHRVCECEGGRYLETEFCIK 120  
 Db 61 LCVPCPDHSYDTSNHTSDCYVCKELQSVKQECNRTNHRVCECEGGRYLETEFCIK 120  
 Qy 121 HRSPPGSGVQAGTPERTNVCCKPCDGFSGETSSKAPCIKHTNCSTFGLLLTQKGNAT 180  
 Db 121 HRSPPGSGVQAGTPERTNVCCKPCDGFSGETSSKAPCIKHTNCSTFGLLLTQKGNAT 180

Qy 181 HDNVCSGNREATQKCGIDVTLCCEAFFRFAVPTKTIIPNWLVLVDSLPCTKVNAESVERI 240  
 Db 181 HDNVCSGNREATQKCGIDVTLCCEAFFRFAVPTKTIIPNWLVLVDSLPCTKVNAESVERI 240  
 Qy 241 KRRHSSQQTQQLLKLWKHKHQRDQEMVKIIQDIDLCSSVQRHLGHSNLTTEQLLALME 300  
 Db 241 KRRHSSQQTQQLLKLWKHKHQRDQEMVKIIQDIDLCSSVQRHLGHSNLTTEQLLALME 300  
 Qy 301 SLPGKKISPEIETRTKCKSEQLLKLLSLWRINKGQDQTLKGLMYALKHLKTSHPFKT 360  
 Db 301 SLPGKKISPEIETRTKCKSEQLLKLLSLWRINKGQDQTLKGLMYALKHLKTSHPFKT 360  
 Qy 361 VTHSLRKTMRFLHSFTMYRLYQKLFLEMGVQSVKISCL 401  
 Db 361 VTHSLRKTMRFLHSFTMYRLYQKLFLEMGVQSVKISCL 401  
 RESULT 2  
 O08727 PRELIMINARY; PRT: 401 AA.  
 AC O08727;  
 DT 01-JUL-1997 (Tremblrel. 04, Created)  
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)  
 DE (OCIF).  
 GN TNFRSF11B OR OPG.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 RN NCBI\_TaxID-10116;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP TISSUE-INTESINE;  
 RX MEDLINE=97262071; PubMed=9108485;  
 RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,  
 Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,  
 Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,  
 Davy E., Bucay N., Renshaw-Gagg L., Hughes T.M., Hill D., Pattison W.,  
 Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,  
 Suggs S., Boyle W.J.;  
 RA "Osteoprotegerin: a novel secreted protein involved in the regulation  
 of bone density.";  
 RL Cell 89:309-319(1997).  
 CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
 OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY  
 SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
 OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
 STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
 CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY  
 SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; U94330; AAB53707.1; -.  
 DR HSSP; P25942; ICDF.  
 DR INTERPRO; IPR001368; -.  
 DR PFAM; PF00020; TNFR\_C6; 4.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; PS0050; TNFR\_NGFR\_2; 2.  
 DR PRODOM; PD000771; -. 1.  
 KW Glycoprotein; Repeat; Cytokine; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 401 OSTEOPROTEGERIN.  
 FT DOMAIN 23 201 4 X TNFR-CYS.  
 FT REPEAT 23 63 TNFR-CYS 1.  
 FT REPEAT 64 106 TNFR-CYS 2.  
 FT REPEAT 107 143 TNFR-CYS 3.  
 FT REPEAT 144 201 TNFR-CYS 4.  
 FT DOMAIN 306 365 DEATH DOMAIN.  
 FT DISULFID 41 54 BY SIMILARITY.  
 FT DISULFID 44 62 BY SIMILARITY.  
 FT DISULFID 65 80 BY SIMILARITY.  
 FT DISULFID 83 97 BY SIMILARITY.



```

FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
SQ SEQUENCE 401 AA; 46192 MW; PEC6A31FD4E573A CRC64;

Query Match 95.4%; Score 2079; DB 11; Length 401;
Best Local Similarity 94.9%; Pred. No. 5e-156;
Matches 379; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 MNKWLCCALLVLDIIEWTTQETLPPKYLHYDPETGHOLLCDKCAPCTYKQHCVTVRKT 60
DB 1 MNKWLCCALLVLDIIEWTTQETLPPKYLHYDPETGHOLLCDKCAPCTYKQHCVTVRKT 60

QY 61 LCVPCPDHSTDSWHTSDCVYSPVKELQSVKQECNRTHNRVCEEGRYLEIEFCLK 120
DB 61 LCVPCPDYSTDSWHTSDCVYSPVKELQSVKQECNRTHNRVCEEGRYLEIEFCLK 120

QY 121 HRCPPGSGVVGAGTPERNVTVCKPCDGFSGFSGETSSKAPCKHTNCSFGLLLIQKGNAT 180
DB 121 HRCPPGLGVLOAGTPERNVTVCKPCDGFSGFSGETSSKAPCKHTNCSFGLLLIQKGNAT 180

QY 181 HDNVCSGNREATOKGIDVTLCFAAFREAVPTKIIPNWSLVLDSPGTVKVAESVERI 240
DB 181 HDNVCSGNREATONCGIDVTLCFAAFREAVPTKIIPNWSLVLDSPGTVKVAESVERI 240

QY 241 KRRHSSQEQTFOLLKWKHQRDQEMVKYTIQDIDLCSSVQRHGLSHNLTTEQLLALME 300
DB 241 KRRHSSQEQTFOLLKWKHQRDQEMVKYTIQDIDLCSSVQRHGLSHNLTTEQLRLIME 300

QY 301 SLPGKKISPEIERTRKTCSSQLLKLKLSLWRKNGDQDTLGLMTALKHLKTSHPPKT 360
DB 301 SLPGKKISPEIERTRKTCSPQLLKLKLSLWRKNGDQDTLGLMTALKHLKRAYHPPKT 360

QY 361 VTHSLRKTIRFLSHFTMYRLYQKLFLEMIGNOVQSVKISCL 401
DB 361 VTHSLRKTIRFLSHFTMYRLYQKLFLEMIGNOVQSVKISCL 401

RESULT 3
ID 000300 PRELIMINARY; PRT; 401 AA.
AC - 000300; 060236;
DT 01-JUL-1997 (TEMBLrel. 04, Created)
DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11B).
GN TNFRSF11B OR OPG OR OCIF.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.O., Wooden S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Suggs S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
RT of bone density."
RL Call 89:309-319(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG FIBROBLAST;

```

```

RX MEDLINE=98151033; PubMed=9492069;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
RA Tsuda E., Morinaga T., Higashio K.,
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
RT osteoclastogenesis in vitro."
RL Endocrinology 139:1329-1337(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=98351569; PubMed=9688283;
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RT "Cloning and characterization of the gene encoding human
RT osteoprotegerin/osteoclastogenesis-inhibitory factor."
RL Eur. J. Biochem. 254:685-691(1998).
CC -/- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -/- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -/- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -/- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.
CC -/- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; AB002146; BAA25910.1; -
DR EMBL; AB008822; BAA32076.1; -
DR EMBL; AB008821; BAA32076.1; JOINED.
DR EMBL; U94332; BAB53709.1; -
DR HSSP; P25942; ICDF.
DR MIM; 602643; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 183 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT CONFLICT 263 263 A -> D (IN REF. 2 AND 3).
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 86.8%; Score 1892; DB 4; Length 401;
Best Local Similarity 85.6%; Pred. No. 2.7e-141;
Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLDI-IEWTTQETLPPKYLHYDPETGHOLLCDKCAPCTYKQHCVTVRK 59
RN [1]
DB 1 MNKLLCCA-LVFLDISIKWTQETFPFKYLHYDEETSHQLLDCRPPGTYLKQHCCTAKWK 59

```

Qy 60 TLVPCPDHSTYDTSWHTSDEVCYSPCKELOSQVKEQNCNRTHNRVCECEGRLYEIEFCL 119  
Db 60 TVCAPDHYTDSWHTSDECLYSPCKELQYVKEQNCNRTHNRVCECKEGRLYEIEFCL 119  
Qy 120 KHRSPGSGVYVAGTPERNIVCKPCPGFSGGETSSKAPCKHTNCSFTGGLLIQKQNA 179  
Db 120 KHRSPGSGVYVAGTPERNIVCKPCPGFSGGETSSKAPCKHTNCSFTGGLLIQKQNA 179  
Qy 180 THDNVCSNREATOKCGIDVTLCCEAFPRFAVPTKIIIPNWSLVLDSPGTVKNAESVER 239  
Db 180 THDNVCSNREATOKCGIDVTLCCEAFPRFAVPTKIIIPNWSLVLDSPGTVKNAESVER 239  
Qy 240 IKRRHSSQEQTFQLLKLKWKHONQDEMVKKIIQDIDLCSSVQRHLGHSNLTTEQLLALM 299  
Db 240 IKRRHSSQEQTFQLLKLKWKHONQDEMVKKIIQDIDLCSSVQRHLGHSNLTTEQLLALM 299  
Qy 300 ESLPCKKISPEIERTKTRCKSSBOLLKLLSLWRKNGDQDTLGLMYALKHLKHTSHPPK 359  
Db 300 ESLPCKKISPEIERTKTRCKSSBOLLKLLSLWRKNGDQDTLGLMYALKHLKHTSHPPK 359  
Qy 360 TVTHSLRKTMRFLHSFTMYRLYQKLFLEMIGNOVQSVKISCL 401  
Db 360 TVTQSLAKTIIRFLHSFTMYRLYQKLFLEMIGNOVQSVKISCL 401

## RESULT 4

Q9UHP4 PRELIMINARY; PRT; 372 AA.  
AC Q9UHP4  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DE OSTEOPROTEGERIN (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;  
RT "Cloning and Expression of Osteoprotegerin from Homo sapiens.";  
RL Sheng Yu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).  
DR EMBL; AF134187; AAF20168.1; -;  
DR HSSP; P25942; 1CDF.  
DR INTERPRO; IPR001368; -;  
DR PFAM; PF00020; TNFR\_C6; 3.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
DR PROSITE; PS50050; TNFR\_NGFR\_2; 2.  
FT NON\_TER 1  
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 82.1%; Score 1789; DB 4; Length 372;  
Best Local Similarity 86.0%; Pred. No. 3,2e-133;  
Matches 320; Conservative 25; Mismatches 27; Indels 0; Gaps 0;

Qy 22 ETLPKYLYHDPETHQLLCKAPGTYLKHQCTVRRKTLVPCPDHSTYDTSWHTSDECV 81  
Db 1 ETLPPKYLYHDEETHQLLCKAPGTYLKHQCTVRRKTLVPCPDHSTYDTSWHTSDECV 60  
Qy 82 YCSPVCKELOSQVKEQNCNRTHNRVCECEGRLYEIEFCLKHSRCPGSGVVOAGTPERNIV 141  
Db 61 YCSPVCKELOSQVKEQNCNRTHNRVCECEGRLYEIEFCLKHSRCPGSGVVOAGTPERNIV 120  
Qy 142 CKKPDGFGFSGTSSKAPCKHTNCSFTGGLLIQKQNAHNVCSNREATOKCGIDVTL 201  
Db 121 CKKPDGFGFSGTSSKAPCKHTNCSFTGGLLIQKQNAHNVCSNREATOKCGIDVTL 180  
Qy 202 CEEAFFRFAVPTKIIPNWSLVLDSPGTVKNAESVERIKRRHSSQEQTFQLLKLKWKHON 261  
Db 181 CEEAFFRFAVPTKIIPNWSLVLDSPGTVKNAESVERIKRRHSSQEQTFQLLKLKWKHON 240

Qy 262 RDQEMVKKIIQDIDLCSSVQRHLGHSNLTTEQLLALMESLPGRKKSIPETIERTKTRCKS 321  
Db 241 KDQDIVKKIIQDIDLCSSVQRHLGHSNLTTEQLLALMESLPGRKKSIPETIERTKTRCKS 300  
Qy 322 SEQLLKLLSLWRKNGDQDTLGLMYALKHLKHTSHPPKTVTHSLRKTMRFLHSFTMYRLY 381  
Db 301 SDQILKLLSLWRKNGDQDTLGLMYALKHLKHTSHPPKTVTHSLRKTMRFLHSFTMYRLY 360  
Qy 382 OKLFLEMIGNOV 393  
Db 361 OKLFLEMIGNOV 372

## RESULT 5

O95407 PRELIMINARY; PRT; 300 AA.  
AC O95407  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).  
GN DCR3 OR TR6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RC MEDLINE=99087326; PubMed=9872321;  
RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,  
RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,  
RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,  
RA Goddard A.D., Botstein D., Ashkenazi A.;  
RT "Genomic amplification of a decoy receptor for Fas ligand in lung and  
RT colon cancer.";  
RL Nature 396:699-703(1998).  
[2]  
RN SEQUENCE FROM N.A.  
RC TISSUE=BLOOD;  
RA Yu K.-Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;  
RT "A newly identified member of tumor necrosis factor receptor  
RT superfamily (TR6) suppresses LIGHT-mediated apoptosis.";  
RL J. Biol. Chem. 274:13733-13736(1999).  
[3]  
RN SEQUENCE FROM N.A.  
RC TISSUE=PANCREAS;  
RA MEDLINE=20122600; PubMed=10655513;  
RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,  
RA Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;  
RT "Overexpression of M68/Dcr3 in human gastrointestinal tract tumors  
RT independent of gene amplification and its location in a four-gene  
RT cluster.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).  
DR EMBL; AF104419; AAD03056.1; -;  
DR EMBL; AF134240; AAD29688.1; -;  
DR EMBL; AF217796; AAF35244.1; -;  
DR EMBL; AF217793; AAF33685.1; -;  
DR EMBL; AF217794; AAF33686.1; -;  
DR HSSP; P25942; 1CDF.  
DR INTERPRO; IPR000561; -;  
DR INTERPRO; IPR001368; -;  
DR PFAM; PF00020; TNFR\_C6; 4.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; UNKNOWN\_1.  
DR PROSITE; PS50050; TNFR\_NGFR\_2; 2.  
DR PRODOM; PD000771; -; 1.  
KW Receptor.  
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 19.5%; Score 424.5; DB 4; Length 300;  
Best Local Similarity 39.0%; Pred. No. 7.1e-26;

Matches 80; Conservative 32; Mismatches 88; Indels 5; Gaps 2;	
QY 26	PXLYHDPETGHOLCDKCAPGYLKHQHTVRRKTLVCPDHSYDTSWHTSDECVYCS 85
Db 34	PTYPWDAETGERLVCAQCPGPFVQPCRRDSPTTCGBCPPPHYTFQFWNLYLRCRYCNV 93
QY 86	VKELQSVKQECNRTNRVCECEGREGYLETEFCLKHRSPPGSGVVOAGTPERTNVCKK 145
Db 94	LCGEREEARACHATHNRACRGTGFFAHAGFCLEHASCPGAGVIAPCTPSQNTOCQC 153
QY 146	PDGFFSGETSSKAPCIKHNTSTFGLLLIQKGNATHDNV---CSGNREATQKCGIDVTLC 202
Db 154	PPGTFSSASSSQCPHNRCTALGLALNVPGSSSHDTLCTSCGTPPLSTRVPGABE--C 211
QY 203	EEAFFRFAVPTKIIPNLSLVDSL 227
Db 212	ERAVIDFAQDISIKRLQRLQAL 236
RESULT 6	
ID Q9PUS0	PRELIMINARY; PRT; 302 AA.
AC Q9PUS0;	
DT 01-MAY-2000	(Tremblrel. 13, Created)
DT 01-MAY-2000	(Tremblrel. 13, Last sequence update)
DT 01-OCT-2000	(Tremblrel. 15, Last annotation update)
DE DECOY RECEPTOR	
OS Salvelinus fontinalis (Brook trout).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;	
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.	
OX NCBI_TaxID=8038;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA Bobe J., Goetz F.W.;	
RT "A tumor necrosis factor receptor homolog is up-regulated in the brook	
RT trout (Salvelinus fontinalis) ovary at the completion of ovulation."	
RL Biol. Reprod. 0:0-0(1999).	
DR EMBL; AF156738; AAD56428.1; -	
DR HSSP; P19438; 1EXT.	
DR INTERPRO; IPR000561; -	
DR INTERPRO; IPR001368; -	
DR PFAM; PF00020; TNFR_C6; 4.	
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.	
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.	
DR PROSITE; PS00050; TNFR_NGFR_2; 1.	
KW Receptor.	
SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;	
Query Match 17.9%; Score 390.5; DB 13; Length 302;	
Best Local Similarity 42.7%; Pred. No. 3.4e-23;	
Matches 67; Conservative 34; Mismatches 53; Indels 3; Gaps 2;	
QY 32	DPETHGOLCDKCAPGYLKHQHTVRRKTLVCPDHSYDTSWHTSDECVYCSVPVCKELQ 91
Db 27	DRYSGLSIVDCPPGTYLAPCSAMKSDCAECPNGAYTEFWNHISKLRCS-MCAENQ 85
QY 92	SVKQECNRTNRVCECEGREGYL--EIEFCLKHRSPPGSGVVOAGTPERTNVCKKCPDGF 149
Db 86	VVKQECSPSNCECECKEGYFNKYEACIKHKECPGPGYANTGTGTPHQDTECVQCAGF 145
QY 150	FSGETSSKAPCIKHNTSTFGLLLIQKGNATHDNVCS 186
Db 146	YSEVSSAKATCLAQSNCKVGLRVVLKQDWHNTLCA 182
RESULT 7	
ID Q62327	PRELIMINARY; PRT; 459 AA.
AC Q62327;	
DT 01-NOV-1996	(Tremblrel. 01, Created)
DT 01-NOV-1996	(Tremblrel. 01, Last sequence update)

DT 01-MAY-2000	(Tremblrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).	
OS Mus musculus (Mouse)	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX NCBI_TaxID=10090;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=NOD;	
RA MEDLINE=95178848; PubMed=7873884;	
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;	
RT "Allelic variation of the type 2 tumor necrosis factor receptor	
RT gene.";	
RL Mamm. Genome 5:726-727(1994).	
DR EMBL; X76401; CAA53981.1; -	
DR HSSP; P19438; 1NCF.	
DR INTERPRO; IPR001368; -	
DR PFAM; PF00020; TNFR_C6; 4.	
DR PROSITE; PS00652; TNFR_NGFR_1; 2.	
DR PROSITE; PS00050; TNFR_NGFR_2; 3.	
FT NON_TER 1	
FT VARIANT 87	S -> T.
FT VARIANT 93	T -> I.
FT VARIANT 268	F -> I.
FT VARIANT 345	S -> F.
FT VARIANT 421	Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;	
Query Match 14.4%; Score 314; DB 11; Length 459;	
Best Local Similarity 34.6%; Pred. No. 5.8e-17;	
Matches 66; Conservative 20; Mismatches 79; Indels 26; Gaps 4;	
QY 18	WTTQETLPPKYL--HYDPEIGH-----QLLCKCAPGYLKHQHTVRRKTLQ 62
Db 2	WATGHTVPAQVLTTPKPEGYEQISQYDRKAQMCACPPGQYVKKHFCNKTSDTVC 61
QY 63	VPQDHSYDTSWHTSDECVYCSVPVCKELQSVKQECNRTNRVCECEGREGYLEIEF---- 117
Db 62	ADCEASMYTQVWNOFRTCLSCSSCSTDQVETRACTQQNRCVCAEAGRYCALTHSGSC 121
QY 118	--CLKHRSPPGSGVVOAGTPERTNVCKKCPDGFSGGETSSKAPCIKHNTSTFGLLLIQ 175
Db 122	RQCMRLSKCGPGFGVASSRAPNGNLCKACAPGTFSDTSTSDVCRPHRCS----ILAI 177
QY 176	KGNATHDNVCS 186
Db 178	PGNASTDAVCA 188
RESULT 8	
ID O88734	PRELIMINARY; PRT; 482 AA.
AC O88734;	
DT 01-NOV-1998	(Tremblrel. 08, Created)
DT 01-NOV-1998	(Tremblrel. 08, Last sequence update)
DT 01-OCT-2000	(Tremblrel. 15, Last annotation update)
DE P80 TNF-ALPHA RECEPTOR.	
GN TNFR2.	
OS Mus musculus (Mouse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX NCBI_TaxID=10090;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA Hurle B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;	
RT "The Mouse Tumor Necrosis Factor Receptor Gene:Genomic Structure and	
RT Characterization of the two Transcripts.";	
RL Genomics 0:0-0(0).	
DR EMBL; Y14619; CAA74969.1; -	
DR EMBL; Y14620; CAA74969.1; JOINED.	
DR EMBL; Y14621; CAA74969.1; JOINED.	
DR EMBL; Y14622; CAA74969.1; JOINED.	





```
FT DISULFID 175 194 BY SIMILARITY.
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 616 AA; 56033 MW; E3DE9A7A08196F81 CRC64;

Query Match 12.2%; Score 266.5; DB 4; Length 616;
Best Local Similarity 24.4%; Pred. No. 4.5e-13;
Matches 108; Conservative 54; Mismatches 166; Indels 115; Gaps 21;

QY 5 LCCALLVLDIETWTOETLPP--KYLHYDPTGHLQCDKCAPGTYLKQHTVRKYLIC 62
DB 16 LLLALLARLOV-----ALQATPTSEKHYE-HLGR--CCNKCEPGKYMSSKCTTSDSVC 68
QY 63 VCPDHSYDTSWHTSDECVYCPVC---KELQSVKQECNTHNRVCECEEGRY--LEIEF 117
DB 69 LFCGDEYLDWNEEDKCL-LHKVCDTGKALVAVVAG-NSTTPRCACVAGVHNSQDEC 126
QY 118 CLKHRSPPGSGVVOAGTPERNTVCKPCDPGFSGTSSKAPCIKHTNCSFGLLLIQKG 177
DB 127 CRNTECAPGLCAQHPLQLNKDCKPCLAGYFSDAFSSDCKRPWTNCTFLGKRVEHGG 186
QY 178 NATHNVCSGNREATQK-----CGDIVTLCFAFRFAVPTKII----- 216
DB 187 TEKSDAVCVSSSLPARKPNPEHYLPGLILL---LFASVALVAIIIFGVCYRKKGKALT 243
QY 217 ---PNWLSVLVDSLPCTKVNAESVERIKRRHSS-----OEOTFOLLKLW 257
DB 244 ANLWHINEACRLSGDK--ESSGSCSVSTHTANFGQQACGCVLLLTLEKTFPEDMCY 301
QY 258 KH-----QNRDQEMVKKIIQDIDLCESVOR-----HLGHSNLTTEQL 295
DB 302 PQDGGVCGTCVGGPGYAGGEDARML-SLVSKTEIEEDSFQMPTEDEYMDRPSQPTDQL 360
QY 296 LALMESLPCKKISP--BEIE-----RTRKT-----KXSSQOLLKLLSLWRK 335
DB 361 LFLTE--FGSKSTPFSEPLEYGENDSLSCQFTGTQSTVGSESCNCETPLCRT----- 411
QY 336 NGDQDTLGLMVALKHLKTSHP 358
DB 412 --DWTPTMSENVLQEVDSGHCP 432

RESULT 13
Q85308 PRELIMINARY; PRT; 355 AA.
AC Q85308;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE SECRETED RECEPTOR BINDING TUMOR NECROSIS FACTOR (CRMB).
GN CRMB.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRIGHTON RED;
RX MEDLINE=83117629; PubMed=6961398;
RA Pickup D.J., Bastia D., Stone H.O., Joklik W.K.;
RT "Sequence of terminal regions of cowpox virus DNA: arrangement of repeated and unique sequence elements."
RL Proc. Natl. Acad. Sci. U.S.A. 79:7112-7116(1982).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRIGHTON RED;
RX MEDLINE=90177240; PubMed=2309453;
RA Parsons B.L., Pickup D.J.;
RT "Transcription of orthopoxvirus telomeres at late times during infection."
RL Virology 175:69-80(1990).
RN [3]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN-BRIGHTON RED;
RX MEDLINE=91196263; PubMed=2014645;
RA Hu F.Q., Pickup D.J.;
RT "Transcription of the terminal loop region of vaccinia virus DNA is initiated from the telomere sequences directing DNA resolution."
RL Virology 181:716-720(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-BRIGHTON RED;
RX MEDLINE=94378510; PubMed=8091665;
RA Hu F.Q., Smith C.A., Pickup D.J.;
RT "Cowpox virus contains two copies of an early gene encoding a soluble secreted form of the type II TNF receptor."
RL Virology 204:343-356(1994).
DR EMBL; L08906; AAA60952.1;
DR HSSP; P19438; IEXT.
DR INTERPRO; IPR001368;
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_2.
DR PROSITE; PS00500; TNFR_NGFR_2; 1.
DR PRODOM; PD000771; 1.
SQ SEQUENCE 355 AA; 39008 MW; 2C9E5C0D42FA4B3 CRC64;

Query Match 10.4%; Score 226; DB 12; Length 355;
Best Local Similarity 24.1%; Pred. No. 3.7e-10;
Matches 85; Conservative 49; Mismatches 141; Indels 78; Gaps 15;

QY 9 LVLVLDIETWTOETLPPKYLHYDPTG-----HQLCDKCAPGTYLKQHC--TVR 57
DB 6 LLLLLLCIIIIISDITP-----HEPSNGKCKDNEYKRHLCLCLSPPGTYASRLCDSKTN 60
QY 58 RKTLCVPCPDHSYDTSWHTSDECVYCPVCVKELQSVKQECNTHNRVCECEGRYLEI-- 115
DB 61 TMTQCTPCASDTFTTRNNHLPACLSGCRGDSNQVETRSCTNTHRICDCAPGYTCFLKG 120
QY 116 ----EFLCKHRSPPGSGVVOAGTPERNTVCKPCDPGFSGTSSKAPCIKHTNCSFG 171
DB 121 SSGCRACVSOFTKCGIGY-GVSCHTPTGVDVCSPCGLGTYSHTVSSVDKC-EPVPSNTFNY 178
QY 172 LLIQKG-NATHDNVC-----SGNREATOKGIDVTL---CEEAFRFAVPTKIIPNMLS 221
DB 179 IDVEINLPVNDTSCRTTTTGLSESISTSELTITMNHKDCDPVERN-----GYFS 229
QY 222 VLVDLSLG-----TKV---NAESVERIKRRHSSQEQTFQLLKLWKHQRDOEM 266
DB 230 VLNEVATSGFTTGQNRQYONISKVCTLNFEIKCNKNDKSYSSSKQ-----LTKKNDOSI 283
QY 267 V--KKIIQDIDLCESVQRHGHSHNLTE-----QLLALMESLPCK 305
DB 284 MPHSESVTLVGDCLSVDIYILYSNTQDYETDITISYHGVNVLVDVDSHMPGR 336

RESULT 14
Q57098 PRELIMINARY; PRT; 349 AA.
AC Q57098;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Camelpox virus (strain CP-1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=28873;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SAUDI-M3;
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87839; AAB94356.1; -.
```

Query Match 10.2%; Score 222.5; DB 12; Length 349;  
Best Local Similarity 25.6%; Pred. No. 6.8e-10;  
Matches 81; Conservative 47; Mismatches 144; Indels 45; Gaps 14;

QY 9 LLVLDDIIEWTQTETLPKYLHYDPTFG-----HQLLCKCAPGTYLKHQC--TVR 57  
|||:||| : ||| : ||| : ||| :  
|||:||| : ||| : ||| : ||| :





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:16:58 ; Search time 70.93 Seconds  
(without alignments)  
193.313 Million cell updates/sec

Title: US-09-389-782A-2

Perfect score: 2179

Sequence: 1 MNKWLCCALLVLDIEWT.....QKLFLEMIGNQSVKISCL 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_36:\*  
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2179	100.0	401	18 W38344	Mouse osteoprotegerin
2	2079	95.4	401	18 W38343	Rat osteoprotegerin
3	1900	87.2	401	20 Y05742	Tumour necrosis fa
4	1900	87.2	401	20 W95030	Tumour necrosis fa
5	1900	87.2	401	20 W83926	Human FTHMA-070 pr
6	1895	87.0	401	17 R99925	Full length osteoc
7	1895	87.0	401	17 W53239	Human OCIF genome
8	1892	86.8	401	18 W38345	Human osteoprotege
9	1892	86.8	401	21 Y43400	Osteoprotegerin pro
10	1891	86.8	401	19 W57635	TR1 receptor prote
11	1885	86.5	401	17 R99932	Mutated OCIF, OCIF
12	1885	86.5	401	17 R99931	Mutated OCIF, OCIF

13	1882	86.4	399	17	R99942	Mutated OCIF, OCIF
14	1880	86.3	401	17	R99934	Mutated OCIF, OCIF
15	1879	86.2	401	17	R99933	Mutated OCIF, OCIF
16	1875	86.0	401	17	R99935	Mutated OCIF, OCIF
17	1862.5	85.5	391	19	W53238	Human OCIF genome
18	1852	85.0	395	19	W57636	Modified TR1 recep
19	1849	84.9	393	17	R99948	Mutated OCIF, OCIF
20	1828	83.9	380	17	R99924	Mature osteoclasto
21	1702.5	78.1	390	17	R99357	Human tumour necro
22	1670.5	76.7	360	17	R99936	Mutated OCIF, OCIF
23	1667	76.5	351	17	R99943	Mutated OCIF, OCIF
24	1645	75.5	359	17	R99939	Mutated OCIF, OCIF
25	1645	75.5	360	17	R99938	Mutated OCIF, OCIF
26	1620	74.3	359	17	R99937	Mutated OCIF, OCIF
27	1539	70.6	327	17	R99941	Mutated OCIF, OCIF
28	1510	69.3	321	17	R99949	Mutated OCIF, OCIF
29	1486.5	68.2	326	17	R99940	Mutated OCIF, OCIF
30	1481.5	68.0	349	20	W83928	Human FTHMA-070 pr
31	1334	61.2	272	17	R99944	Mutated OCIF, OCIF
32	1096	50.3	420	20	W89224	Tumour necrosis fa
33	1075.5	49.4	417	20	W89226	Tumour necrosis fa
34	1073	49.2	208	20	W89231	Mouse osteoprotege
35	1029	47.2	208	20	W89230	Rat osteoprotegeri
36	973	44.7	197	17	R99945	Mutated OCIF, OCIF
37	971.5	44.6	397	20	W89227	Tumour necrosis fa
38	946	43.4	208	20	W89232	Human osteoprotege
39	933	42.8	187	21	Y77464	Primate protein
40	928	42.6	187	17	R99950	Mutated OCIF, OCIF
41	810	37.2	366	20	W89228	Tumour necrosis fa
42	709	32.5	143	17	R99946	Mutated OCIF, OCIF
43	666.5	30.6	145	17	R99930	Osteoclastogenesis
44	665	30.5	154	17	R99929	Osteoclastogenesis
45	536.5	24.6	311	20	W89229	Tumour necrosis fa

#### ALIGNMENTS

RESULT	1
W38344	W38344 standard; Protein; 401 AA.
ID	XX
AC	XX
W38344;	XX
DT	20-APR-1998 (first entry)
XX	Mouse osteoprotegerin.
DE	XX
XX	Osteoprotegerin; antibody; diagnosis; affinity purification;
KW	recombinant production; transgenic animal; treatment; prevention;
KW	antisense oligonucleotide; probe; detection; screening; mouse;
KW	bone disease; osteoporosis; Paget's disease; hypercalcaemia;
KW	hyperparathyroidism; rheumatoid arthritis; osteomyelitis;
KW	osteolytic metastasis; periodontal bone loss; bone necrosis;
XX	osteopaenia; murine.
OS	Mus sp.
XX	XX
Key	Location/Qualifiers
CDS	90..1295
FT	/*tag= a
FT	/product= osteoprotegerin
DE	DE19654610-A1.
XX	26-JUN-1997.
XX	20-DEC-1996; 96DE-1054610.
XX	03-SEP-1996; 96US-0706945.
PR	22-DEC-1995; 95US-0577788.
XX	(AMGE-) AMGEN INC.
PA	

XX Boyle WJ, Calzone FJ, Lacey DL, Chang MS;  
XX WPI: 1997-334271/31.  
DR N-PSDB: T96062.  
XX Nucleic acid encoding osteoprotegerin - useful for treatment of  
PT diseases involving excessive bone loss, e.g. osteoporosis  
XX  
XX Claim 23; Pages 106-107; 182pp; German.  
XX  
XX The present sequence is mouse osteoprotegerin (OPG). Anti-OPG  
CC antibodies can be used in OPG diagnostic assays, and as affinity  
CC purification materials. The OPG cDNA can be used to express  
CC recombinant OPG and to generate transgenic animals. It can also  
CC be used to regulate the level of OPG in mammals, specifically to  
CC increase OPG levels, however the use of antisense sequences is  
CC also contemplated. Fragments of the cDNA can be used as probes to  
CC detect OPG expressing cells and tissue, and to screen cDNA  
CC libraries for related sequences. OPG can be used to treat or  
CC prevent bone diseases, specifically excessive bone loss, e.g.  
CC osteoporosis, Paget's disease, hypercalcaemia,  
CC hyperparathyroidism, rheumatoid arthritis, osteomyelitis, and  
CC osteolytic metastases, periodontal bone loss, bone necrosis and  
XX osteopaenia.  
XX  
XX Sequence 401 AA;  
XX  
Query Match 100.0%; Score 2179; DB 18; Length 401;  
Best Local Similarity 100.0%; Pred. No. 2.le-179;  
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNKWLCCALLVLLDIETWTQETLPKYLHYDPTGHLCDKCAPGYLKQHTVRKKT 60  
Db 1 mnkwlccallvllldietwtqetlppkylhydpetghlclcdkcapgytlkqhtvrkt 60  
QY 61 LCVPDHSYTDSDWHTSDECVYCSVKELQSVKQECNRTHNRVCEEGRYLEIEFCLK 120  
Db 61 lcvpdhsytdsdwhtsdecvycspvckelqsvkqecnrthnrviceeegryleiefclk 120  
QY 121 HRSPGSGVQAGTPERTVCKKCPDGFSGFTSSKAPCIKHTNCSTFGLLLLQKGNAT 180  
Db 121 hrscpgsgvqagtpertvckkcpdgffsgftsskapcikhntcstfgllllqkgnat 180  
QY 181 HDNVCSGNREATQKCGIDVTLCCEAFFRAVPTKIIIPNWLVSVDLPSTKVNESVERI 240  
Db 181 hdnvcsgnreatqkcgidvtlceaffravptkiiipnwlsvdlsipgtkvnasveri 240  
QY 241 KRRHSSQEQTFOLLKWKHQRDQEMVKIIQDIDLCESSYORHLGHSNLTTEQLALME 300  
Db 241 krrhssqeqtfllklwkhqrdqemvkkiiqdidlcessygrhlghsnltteqlalme 300  
QY 301 SLPGKKISPEIERTRKTKCSSEBOLLKLLSWRIKNGDQDTLGLMVALKHLKTSHPKPT 360  
Db 301 slpgkkspeiertrktkcsseqllkllslwrikngdqdtlglmyalkhlktshtpkt 360  
QY 361 VTHSLRKMRFTHSFMTYRLYKLFLEMGVQSVKISCL 401  
Db 361 vthslrkmrftmrylqkflflemgqvsvkislcl 401  
RESULT 2  
ID W38343  
XX W38343 standard; Protein; 401 AA.  
AC W38343;  
XX  
DT 20-APR-1998 (first entry)  
XX  
DE Rat osteoprotegerin.  
XX  
KW Rat; osteoprotegerin; antibody; diagnosis; affinity purification;

KW recombinant production; transgenic animal; treatment; prevention;  
KW antisense oligonucleotide; probe; detection; screening;  
KW bone disease; osteoporosis; Paget's disease; hypercalcaemia;  
KW hyperparathyroidism; rheumatoid arthritis; osteomyelitis;  
KW osteolytic metastasis; periodontal bone loss; bone necrosis;  
KW osteopaenia.  
XX Rattus sp.  
XX OS  
XX DE19654610-A1.  
XX  
XX 26-JUN-1997.  
XX  
XX 20-DEC-1996; 96DE-1054610.  
XX  
XX 03-SEP-1996; 96US-0706945.  
XX  
XX 22-DEC-1995; 95US-0577788.  
XX  
XX (AMGE-) AMGEN INC.  
XX  
XX Boyle WJ, Calzone FJ, Lacey DL, Chang MS;  
XX WPI: 1997-334271/31.  
DR N-PSDB: T96061.  
XX  
XX Nucleic acid encoding osteoprotegerin - useful for treatment of  
PT diseases involving excessive bone loss, e.g. osteoporosis  
XX  
XX Claim 23; Pages 102-104; 182pp; German.  
XX  
XX The present sequence is rat osteoprotegerin (OPG). Anti-OPG  
CC antibodies can be used in OPG diagnostic assays, and as affinity  
CC purification materials. The OPG cDNA can be used to express  
CC recombinant OPG and to generate transgenic animals. It can also  
CC be used to regulate the level of OPG in mammals, specifically to  
CC increase OPG levels, however the use of antisense sequences is  
CC also contemplated. Fragments of the cDNA can be used as probes to  
CC detect OPG expressing cells and tissue, and to screen cDNA  
CC libraries for related sequences. OPG can be used to treat or  
CC prevent bone diseases, specifically excessive bone loss, e.g.  
CC osteoporosis, Paget's disease, hypercalcaemia,  
CC hyperparathyroidism, rheumatoid arthritis, osteomyelitis, and  
CC osteolytic metastases, periodontal bone loss, bone necrosis and  
XX osteopaenia.  
XX  
XX Sequence 401 AA;  
XX  
Query Match 95.4%; Score 2079; DB 18; Length 401;  
Best Local Similarity 94.5%; Pred. No. 8.5e-171;  
Matches 379; Conservative 10; Mismatches 12; Indels 0; Gaps 0;  
QY 1 MNKWLCCALLVLLDIETWTQETLPKYLHYDPTGHLCDKCAPGYLKQHTVRKKT 60  
Db 1 mnkwlccallvllldietwtqetlppkylhydpetghlclcdkcapgytlkqhtvrkt 60  
QY 61 LCVPDHSYTDSDWHTSDECVYCSVKELQSVKQECNRTHNRVCEEGRYLEIEFCLK 120  
Db 61 lcvpdhsytdsdwhtsdecvycspvckelqsvkqecnrthnrviceeegryleiefclk 120  
QY 121 HRSPGSGVQAGTPERTVCKKCPDGFSGFTSSKAPCIKHTNCSTFGLLLLQKGNAT 180  
Db 121 hrscpgsgvqagtpertvckkcpdgffsgftsskapcikhntcstfgllllqkgnat 180  
QY 181 HDNVCSGNREATQKCGIDVTLCCEAFFRAVPTKIIIPNWLVSVDLPSTKVNESVERI 240  
Db 181 hdnvcsgnreatqkcgidvtlceaffravptkiiipnwlsvdlsipgtkvnasveri 240  
QY 241 KRRHSSQEQTFOLLKWKHQRDQEMVKIIQDIDLCESSYORHLGHSNLTTEQLALME 300  
Db 241 krrhssqeqtfllklwkhqrdqemvkkiiqdidlcessygrhlghsnltteqlalme 300  
QY 301 SLPGKKISPEIERTRKTKCSSEBOLLKLLSWRIKNGDQDTLGLMVALKHLKTSHPKPT 360

Db 301 slpgkispdeiertrkckpseqlklislrwknkgdgtlkgymalkhlykayhfpkt 360  
 QY 361 vthslrktwrfelshfmyrlyoklflemignovqsvkiscsl 401  
 Db 361 vthslrktirflshfmyrlyoklflemignovqsvkiscsl 401

RESULT 3

Y05742  
 ID Y05742 standard; Protein; 401 AA.  
 AC Y05742;  
 DT 19-JUL-1999 (first entry)  
 XX Tumour necrosis factor receptor TRL.  
 DE Tumour necrosis factor receptor; TRL; osteoprotegerin; agonist;  
 KW antagonist; screening; human; cancer; AIDS; Alzheimer's disease;  
 KW inflammation; arthritis; septicemia; autoimmune disease;  
 KW psoriasis; inflammatory bowel disease; transplant rejection;  
 KW graft versus host disease; infection; stroke; ischemia;  
 KW acute respiratory disease syndrome; restenosis; brain injury;  
 KW bone disease; atherosclerosis; therapy.  
 OS Homo sapiens.  
 XX  
 PN EP911633-A1.  
 XX  
 PD 28-APR-1999.  
 XX  
 PF 02-OCT-1998; 98EP-0203332.  
 XX  
 PR 08-OCT-1997; 97US-0061334.  
 XX  
 PA (SMK ) SMITHKLINE BEECHAM CORP.  
 XX  
 PI McDonnell PC, Young PR, Zou J;  
 XX  
 DR WPI; 1999-246560/21.  
 XX  
 PT Identifying agonists and antagonists of tumor necrosis factor  
 PT related receptors TRL, TR3 and TR5, and of ligand TL3, useful for  
 PT treatment of cancer, AIDS, Alzheimer's disease, bone disease etc  
 XX  
 PS Disclosure; Page 10-12; 23pp; English.

XX The present sequence represents tumour necrosis factor receptor  
 CC (TNFR) TRL, also known as osteoprotegerin. The invention relates  
 CC to TNFR related polypeptides TRL, TR3 and TR5 (see Y05742-44) and  
 CC their ligand TL3 (see Y05745). TRL, TR3, TR5 and TL3 are used in  
 CC claimed methods of identifying agonists and antagonists, i.e.  
 CC compounds that bind to the receptors or ligand, and which activate  
 CC (agonist) or inhibit activation of (antagonists) TRL, TR3, TR5 or  
 CC TL3. A screening kit for identifying agonists, antagonists,  
 CC ligands, receptors, substrates, enzymes etc. for TRL, TR3, TR5 or  
 CC TL3 polypeptides is provided. The agonists and antagonists are  
 CC useful for treatment of chronic and acute inflammation, arthritis,  
 CC septicemia, autoimmune disease e.g. inflammatory bowel disease,  
 CC psoriasis, transplant rejection, graft versus host disease,  
 CC infection, stroke, ischemia, acute respiratory disease syndrome,  
 CC restenosis, brain injury, AIDS, bone diseases, cancer (e.g.  
 CC lymphoproliferative disorders), atherosclerosis and Alzheimer's  
 CC disease, etc., caused by imbalance of TRL, TR3, TR5 or TL3.

XX Sequence 401 AA;

Query Match 87.2%; Score 1900; DB 20; Length 401;  
 Best Local Similarity 85.8%; Pred. No. 2.1e-155;  
 Matches 345; Conservative 26; Mismatches 29; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLDI-I EWTTQETLPPKYLHYDPETGHQLLCKAPGYLKHQCTVRRK 59  
 Db 1 mnlkllcca-lvfidisikwtgtetfpkylhydeetshqllcdkcpptylkqhctakwk 59  
 QY 60 TLCVPCPDHSYTSWHTSDCVYCSVKELQSVKQECNTHNRVCECEGRYLEIEFCL 119  
 Db 60 tvcapcpdhyttdswhtsdeclycspvckelgyvkqecnrthrvceckeegyleiefcl 119  
 QY 120 KHRSCPPGGVVOAGTPERNTVCKKCPDGFSGSETSSKAPCIKHTNCSTFGLLLIQKNA 179  
 Db 120 khrcppggvgvqagtperntvckrcpdgffsnetsskaperkhtncsvfgllltqkna 179  
 QY 180 THDNVSGNREATQKCGIDVTLCCEAFFRFAVPTKIIPNWLSDVSLPCTKVNAESVER 239  
 Db 180 thdnicsgnsestqcgidvtlceaffrfavptkftpnwslvldnlpotkvnaesver 239  
 QY 240 IKRRHSSQEQTFOLLKLWKHQNDRQEMVKIIQDIDLCSSVORHLGHSNLTTEQLLALM 299  
 Db 240 ikrrhssqeqtfqllklwkhnkndqdvkiiqdidlcensvgrhghantfeqlrslm 299  
 QY 300 ESLPGKKISPEEIERTRKCKSSFOALLKLWLRIKNGDQDTLGLMYALKHLKTSHPK 359  
 Db 300 eslpgkksqeqtfqllklwkhnkndqdvkiiqdidlcensvgrhghantfeqlrslm 359  
 QY 360 TVTSHLRKTRMRFLHSETMYRYOKLFLEMIGNOVQSVKISCL 401  
 Db 360 tvtqslktrmrflhsetmyryoklflemignovqsvkiscsl 401  
 RESULT 4  
 ID - W95030 standard; protein; 401 AA.  
 AC W95030;  
 DT 13-MAY-1999 (first entry)  
 XX Tumour necrosis factor receptor (TNF-R) related polypeptide TRL.  
 DE Tumour necrosis factor receptor; TNF-R; TRL; TR2; TL2; arthritis;  
 KW inflammation; septicemia; autoimmune disease; transplant rejection;  
 KW graft vs. host disease; infection; stroke; ischemia; brain injury; AIDS;  
 KW acute respiratory disease syndrome; restenosis; bone disease; cancer;  
 KW atherosclerosis; Alzheimer's disease.  
 XX  
 OS Unidentified.  
 XX  
 PN EP897114-A2.  
 XX  
 PD 17-FEB-1999.  
 XX  
 PF 04-JUN-1998; 98EP-0304424.  
 XX  
 PR 29-AUG-1997; 97US-0057550.  
 PR 13-AUG-1997; 97US-0055513.  
 PR 26-AUG-1997; 97US-0056980.  
 XX (SMK ) SMITHKLINE BEECHAM CORP.  
 XX  
 PI Brigham-burke MR, Young PR;  
 XX WPI; 1999-134308/12.  
 DR  
 XX Identifying agonists and antagonists to tumour necrosis factor  
 PT receptor (TNF-R) related polypeptides (TRL, TR2, TL2 and TL4) -  
 PT useful for treating stroke, Alzheimer's disease and AIDS  
 XX  
 PS Disclosure; Page 11-12; 18pp; English.  
 XX  
 CC The invention relates to identifying agonists or antagonists to tumour  
 CC necrosis factor receptor (TNF-R) related polypeptides (TRL and TR2, TL2  
 CC and TL4). The method comprises: (a) (i) contacting TRL or TR2 with a  
 CC candidate compound in the presence of TL2 or TL4; or (ii) contacting TL2

CC or TL4 with a candidate compound in the presence of TR1 or TR2; and (b)  
 CC assessing the ability of the candidate compound to compete with TR1 or  
 CC TR2 binding to TL4. TR and TL agonists and antagonists are useful  
 CC for treating diseases caused by imbalance of TL or TR polypeptide levels,  
 CC which cause: chronic and acute inflammation, arthritis, septicemia,  
 CC autoimmune diseases, transplant rejection, graft vs. host disease,  
 CC infection, stroke, ischemia, acute respiratory disease syndrome,  
 CC restenosis, brain injury, AIDS, bone diseases, cancer, atherosclerosis  
 CC and Alzheimer's disease. The present sequence represents a TNF-R related  
 CC polypeptide TR1.  
 XX  
 SQ Sequence 401 AA;

Query Match 87.2%; Score 1900; DB 20; Length 401;  
 Best Local Similarity 85.8%; Pred. No. 2.1e-155;  
 Matches 345; Conservative 26; Mismatches 29; Indels 2; Gaps 2;  
 QY 1 MNKWLCCALLVLLDI-IETWTTQETLPPKYLHYDPETGHQLLCDRCAPGTYLKQHCTVRRK 59  
 DB 1 mnkllcca-lvldisikwtqetfppkylhydeetshqllcdkcpqgtylkqhtakwk 59  
 QY 60 TLCVPCPDHSTDSWHTSDECVYSPVCKELQSVKQECNTHNRVCEEGRYLEIEFCL 119  
 DB 60 tvcapcpdhyttdswhtsdeclycspvckelqyvkqecnrthnrvcceckryleiefc 119  
 QY 120 KHRSCPPGSGVVQAGTPTERNVCKKCPDGPFGSTSSKAPCIKHTNCSTFGLLLIQKNA 179  
 DB 120 khrcscppgvgvqagtperntvckrcpdgffsnetskapcrkhtncsvfllltqkqna 179  
 QY 180 THDNVCSGNREATOKCGIDVTLCCEAFRRFAVPTKIIPNWLVSILVDSLPGTKVNAESVER 239  
 DB 180 thdnicsgnsesctqcgldvtlceeafrfavpckftpnwlsivdnlpgtkvnaesver 239  
 QY 240 IKRRHSSQEQTFOLLKWLKHONRDOEMVKTIQDIDICESVQRHLGHSNLTFTQLALM 299  
 DB 240 ikrhsssqeqtfllklwkhnkdqdivkklididicensvqrhghantfeqlrslm 299  
 QY 300 ESLPGKISPEIERTRKTKSSQOLKLLSLWRKNGDQDTLGLMYALKHLKTSHPFK 359  
 DB 300 eslpgkkgvaediektikackpsdqilklslwrkngdqdtklglmhalkhsktyhfpk 359  
 QY 360 TVTSLRKTMRFLHSFTMYRLYQKLFLEMIGNQVQSVKISCL 401  
 DB 360 tvtslkkktirflhsftmyklyqlflemignqvsvkisc 401

RESULT 5  
 W83926  
 ID W83926 standard; Protein; 401 AA.  
 AC AC  
 AC W83926;  
 XX  
 DT 01-MAR-1999 (first entry)  
 XX Human FTHMA-070 protein.  
 DE FTHMA-070; human; neurological disorder; therapy; diagnosis.  
 KW Homo sapiens.  
 XX  
 OS  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /label= Sig\_peptide  
 FT Protein 22..401  
 FT /label= Mat\_protein  
 XX  
 FN W09848051-A2.  
 XX  
 PD 29-OCT-1998.  
 XX  
 PF 17-APR-1998; 98WO-US07714.  
 XX

PR 10-OCT-1997; 97US-0062017.  
 PR 18-APR-1997; 97US-0044746.  
 XX  
 PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.  
 XX  
 PI Holtzman D, McCarthy SA;  
 XX  
 DR WPI; 1999-024021/02.  
 DR N-PSDB; V69277.  
 XX  
 PT New isolated human FTHMA-070 and T85 proteins - used to develop  
 PT products for the diagnosis and therapy of disorders involving  
 PT cellular processes, e.g. neuronal development.  
 XX  
 PS Claim 8; Fig 1; 127pp; English.  
 XX  
 CC This is the amino acid sequence of human FTHMA-070, a novel protein  
 CC having homology to tumour necrosis factor receptor. The sequence  
 CC was deduced from that of a cDNA clone (see V69277) isolated from a  
 CC cardiac coronary artery smooth muscle cell library. FTHMA-070  
 CC nucleic acids and polypeptides of the invention are useful as  
 CC modulating agents in regulating a variety of cellular processes.  
 CC They can be used for identifying compounds which bind to or modulate  
 CC the activity of the polypeptides (claimed). They can also be used  
 CC in screening assays, detection assays (e.g. chromosomal mapping,  
 CC tissue typing, forensic biology), predictive medicine (e.g.  
 CC diagnostic assays, prognostic assays, monitoring clinical trials,  
 CC and pharmacogenomics), and methods of treatment (e.g. therapeutic  
 CC and prophylactic) e.g. for neurological disorders.  
 XX  
 SQ Sequence 401 AA;

Query Match 87.2%; Score 1900; DB 20; Length 401;  
 Best Local Similarity 85.8%; Pred. No. 2.1e-155;  
 Matches 345; Conservative 26; Mismatches 29; Indels 2; Gaps 2;  
 QY 1 MNKWLCCALLVLLDI-IETWTTQETLPPKYLHYDPETGHQLLCDRCAPGTYLKQHCTVRRK 59  
 DB 1 mnkllcca-lvldisikwtqetfppkylhydeetshqllcdkcpqgtylkqhtakwk 59  
 QY 60 TLCVPCPDHSTDSWHTSDECVYSPVCKELQSVKQECNTHNRVCEEGRYLEIEFCL 119  
 DB 60 tvcapcpdhyttdswhtsdeclycspvckelqyvkqecnrthnrvcceckryleiefc 119  
 QY 120 KHRSCPPGSGVVQAGTPTERNVCKKCPDGPFGSTSSKAPCIKHTNCSTFGLLLIQKNA 179  
 DB 120 khrcscppgvgvqagtperntvckrcpdgffsnetskapcrkhtncsvfllltqkqna 179  
 QY 180 THDNVCSGNREATOKCGIDVTLCCEAFRRFAVPTKIIPNWLVSILVDSLPGTKVNAESVER 239  
 DB 180 thdnicsgnsesctqcgldvtlceeafrfavpckftpnwlsivdnlpgtkvnaesver 239  
 QY 240 IKRRHSSQEQTFOLLKWLKHONRDOEMVKTIQDIDICESVQRHLGHSNLTFTQLALM 299  
 DB 240 ikrhsssqeqtfllklwkhnkdqdivkklididicensvqrhghantfeqlrslm 299  
 QY 300 ESLPGKISPEIERTRKTKSSQOLKLLSLWRKNGDQDTLGLMYALKHLKTSHPFK 359  
 DB 300 eslpgkkgvaediektikackpsdqilklslwrkngdqdtklglmhalkhsktyhfpk 359  
 QY 360 TVTSLRKTMRFLHSFTMYRLYQKLFLEMIGNQVQSVKISCL 401  
 DB 360 tvtslkkktirflhsftmyklyqlflemignqvsvkisc 401

RESULT 6  
 R99925  
 ID R99925 standard; Protein; 401 AA.  
 XX  
 AC R99925;  
 XX  
 DT 22-APR-1997 (first entry)

XX Full length osteoclastogenesis inhibitory factor.  
 XX Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW Osteoporosis.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..401  
 FT /note= "Mature OCIF, claim 6"  
 XX  
 XX W09626217-A1.  
 XX  
 XX 29-AUG-1996.  
 XX  
 XX 20-FEB-1996; 96WO-JP00374.  
 XX  
 XX 21-JUL-1995; 95JP-0207508.  
 PR 20-FEB-1995; 95JP-0054977.  
 XX  
 XX (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PA  
 XX Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 XX  
 XX WPI; 1996-402320/40.  
 DR N-PSDB; T36685.  
 XX  
 XX DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 XX  
 XX Disclosure; Page 64-66; 183pp; Japanese.  
 XX  
 XX This sequence represents the full length osteoclastogenesis inhibitory  
 CC factor (OCIF) of the invention. The OCIF has a molecular weight by  
 CC SDS-PAGE of 60 kD under reducing conditions and 120 kD under non-  
 CC reducing conditions. The protein is adsorbed onto cation-exchangers  
 CC or heparin and its activity is lowered after 10 mins at 70 deg.C or  
 CC 30 mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is  
 CC useful in the control of bone resorption and therefore in the  
 CC treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 XX  
 XX Sequence 401 AA;  
 SQ  
 Query Match 87.0%; Score 1895; DB 17; Length 401;  
 Best Local Similarity 85.6%; Pred. No. 5.6e-155;  
 Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;  
 QY 1 MNKWLCCALLVLDI-TEWTTQETFLPKYLYHPDETHQGLLDCDKAPGTYLKQHCVTYRRK 59  
 Db 1 mnnllcca-lvfidisikwtetfpkylhydeetshqllcdkcpptgylkqhtakwk 59  
 QY 60 TLCVPCPDHSYTDSWHTSDSCVPCVKELQSVKQECNTHNRVCEBGRYLEIEFCL 119  
 Db 60 tvcaepcdpdytdswhtsdsclycspvckelqyvkqecnthnrvcckegyleiefcl 119  
 QY 120 KHRSCPPGSGVVOAGTERTNVCKCPDGFSGTSSKAPCIKHTNCTSTGLLLIQKGN 179  
 Db 120 khrcpppgfvvgaqtpertntvkcrcpdgffsnetsskpcrkhtncsvfllltqkgn 179  
 QY 180 THDNVCSGNREATQCGIDVTLCEAEFFRAVPTKIIPNMLSLVLDLSLPGTKVNAESVER 239  
 Db 180 thdnicsgsestqcgldvtlceaeaffravptkftpnwlvldnlpgtkvnaesver 239  
 QY 240 IKRRHSSQEQTFQLLKLWKNQRDQEWVKLIQIDICESSVQRHLGHSNLTTEQLALM 299  
 Db 240 ikrdhssgeqtfqllklwkhqndqdivkklqgidicessvqgrhlanltfegrlslm 299

QY 300 ESIPGKKISPEIERTKTKCSSEQLKLKLSLWRKIKNGDQDTLKLGLMYALKHLKTSHPK 359  
 Db 300 esipgkkvgaediektikackpsdqllklslwrkngdqtikglmhalkhsktyhfpk 359  
 QY 360 TVVHSLRKTRMRFILHSFTMYRLYOKLFLEMTGNQVQSVKISCL 401  
 Db 360 tvtqslkktirflhsftmyklyqklflemignqvgsvkiscsl 401  
 RESULT 7  
 W53239  
 ID W53239 standard; Protein; 401 AA.  
 XX  
 AC W53239;  
 XX  
 DT 15-JUL-1998 (first entry)  
 XX  
 DE Human OCIF genome DNA protein.  
 XX  
 KW Human; OCIF; genome; osteoclast; antipyretic; osteoporosis;  
 KW rheumatism; multiple sclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /label= signal  
 FT Protein 22..401  
 FT /label= OCIF\_protein  
 XX  
 XX W09807840-A1.  
 PN  
 XX 26-FEB-1998.  
 PD  
 XX 19-AUG-1997; 97WO-JP02859.  
 PF  
 XX 19-AUG-1996; 96JP-0235928.  
 PR  
 XX (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 PA  
 XX Morinaga T, Nakagawa N, Yasuda H;  
 PI  
 XX WPI; 1998-169150/45.  
 DR N-PSDB; V20768.  
 XX  
 XX Inhibition of osteoclast formation and/or antipyretic activity -  
 PT useful for, e.g. treating osteoporosis, rheumatism and multiple  
 PT sclerosis  
 XX  
 XX Claim 4; Page 26-28; 36pp; English.  
 PS  
 XX The present sequence represents human OCIF genome DNA protein which is  
 CC specifically claimed in the present invention. The present invention  
 CC describes: (i) a method of inhibiting the formation of osteoclasts  
 CC and/or antipyretic proteins, which have the following characteristics:  
 CC (i) MW determined by SDS-PAGE of approximately 60 kDa under reducing  
 CC conditions, and 60 kDa and 120 kDa under non-reducing conditions; (ii)  
 CC 401 aa sequence (III) (see W53239); (iii) affinity to cation exchangers  
 CC and heparin; (iv) its osteoclast formation inhibiting activity reduced  
 CC by heating at 70 degrees Celsius for 10 minutes or 56 degrees Celsius  
 CC for 30 minutes, and (v) its antipyretic activity reduced by heating at  
 CC 90 degrees Celsius for 10 minutes, and (2) a method for preparing the  
 CC above proteins. The proteins are useful for, e.g. treatment and  
 CC prevention of osteoporosis, rheumatism or multiple sclerosis, and also  
 CC as antigens for immunological diagnosis of these diseases and disorders.  
 XX  
 XX Sequence 401 AA;

Query Match 87.0%; Score 1895; DB 19; Length 401;  
 Best Local Similarity 85.6%; Pred. No. 5.6e-155;  
 Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;

Qy 1 MNKWLCCALLVLLDI-IEWTTQETLPPKYLHYDPETGHQLLDCDCAPGTYLKQHCTVRRK 59  
 Db 1 mnnllcca-lvfidisikwtqetfppkylhydeetsqllcdkcpptylkqhtakwk 59  
 Qy 60 TLCVPCPDHSTYDTSWHTSDECVYCSVCKELQSVKQECNRTHNRVCEEGRYLEIEFCL 119  
 Db 60 tvacpcpdhytydswhtsdeclycspvckelqyvkqecnrthnrveckeegryleiefcl 119  
 Qy 120 KHRSCPPGSGVVGAGTPTERNVCKKCPDGFSGTSSKAPCIKHTNCSFTGILLIQKNA 179  
 Db 120 khrcppgfvvgagtpertntvckrcpdgffsnetskapcrkhtncsvfgllltqkna 179  
 Qy 180 THDNVCSGNREATOKCGIDVTICEAFAFRFAVPTKIIPNWLVSVLVDSPGTVKNAESVER 239  
 Db 180 thdnicsgnsestqkcgidvticeaaffrfaftpwnlsvldnlpgtkvnaesver 239  
 Qy 240 IKRRHSQEQTFQLLKLWKNQRDQEMVKYKIIQDIDCESSVQRHLGHSNLTTEQLLALM 299  
 Db 240 ikrhssqeqtfqlklwkhnqkdqkvkiiqdidlccensvqrhghanitfeqlrslm 299  
 Qy 300 ESLPGKKISPEIERTRKTCSSQOLKLLSLWRKNGDQDTLGLMYALKHLSHFPK 359  
 Db 300 eslpkgkvgaediektikackpsdqllkllslwrkngdgtlkglmhalkhsktyhfpk 359  
 Qy 360 TVTHSLRKTMRFLHSFTMYRYQKLFLEMIGNOVQSVKISCL 401  
 Db 360 tvtsqlkktirflhsftmykyqlflemignqvsvkiscsl 401

RESULT 8  
 W38345  
 ID W38345 standard; Protein; 401 AA.  
 XX W38345;  
 AC W38345;  
 DT 20-APR-1998 (first entry)  
 DE Human osteoprotegerin.  
 KW Osteoprotegerin; antibody; diagnosis; diagnosis; affinity purification;  
 KW recombinant production; transgenic animal; treatment; prevention;  
 KW antisense oligonucleotide; probe; detection; screening; human;  
 KW bone disease; osteoporosis; Paget's disease; hypercalcaemia;  
 KW hyperparathyroidism; rheumatoid arthritis; osteomyelitis;  
 KW osteolytic metastasis; periodontal bone loss; bone necrosis;  
 KW osteopaenia.  
 XX Homo sapiens.  
 XX DE19654610-A1.  
 XX 26-JUN-1997.  
 XX 20-DEC-1996; 96DE-1054610.  
 XX 03-SEP-1996; 96US-0706945.  
 XX 22-DEC-1995; 95US-0577788.  
 XX (AMGE-) AMGEN INC.  
 XX Boyle WJ, Calzone FJ, Lacey DL, Chang MS;  
 XX WPI; 1997-334271/31.  
 XX N-FSDB; T96063.  
 XX Nucleic acid encoding osteoprotegerin - useful for treatment of  
 XX diseases involving excessive bone loss, e.g. osteoporosis  
 XX Claim 23; Pages 109-111; 182pp; German.  
 XX The present sequence is human osteoprotegerin (OPG). Anti-OPG  
 XX antibodies can be used in OPG diagnostic assays, and as affinity  
 XX purification materials. The OPG cDNA can be used to express

CC recombinant OPG and to generate transgenic animals. It can also  
 CC be used to regulate the level of OPG in mammals, specifically to  
 CC increase OPG levels, however the use of antisense sequences is  
 CC also contemplated. Fragments of the cDNA can be used as probes to  
 CC detect OPG expressing cells and tissue, and to screen cDNA  
 CC libraries for related sequences. OPG can be used to treat or  
 CC prevent bone diseases, specifically excessive bone loss, e.g.  
 CC osteoporosis, Paget's disease, hypercalcaemia,  
 CC hyperparathyroidism, rheumatoid arthritis, osteomyelitis,  
 CC osteolytic metastases, periodontal bone loss, bone necrosis and  
 CC osteopaenia.  
 XX  
 XX Sequence 401 AA;  
 Qy  
 Query Match 86.8%; Score 1892; DB 18; Length 401;  
 Best Local Similarity 85.6%; Pred. No. le-154;  
 Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;  
 Qy 1 MNKWLCCALLVLLDI-IEWTTQETLPPKYLHYDPETGHQLLDCDCAPGTYLKQHCTVRRK 59  
 Db 1 mnnllcca-lvfidisikwtqetfppkylhydeetsqllcdkcpptylkqhtakwk 59  
 Qy 60 TLCVPCPDHSTYDTSWHTSDECVYCSVCKELQSVKQECNRTHNRVCEEGRYLEIEFCL 119  
 Db 60 tvacpcpdhytydswhtsdeclycspvckelqyvkqecnrthnrveckeegryleiefcl 119  
 Qy 120 KHRSCPPGSGVVGAGTPTERNVCKKCPDGFSGTSSKAPCIKHTNCSFTGILLIQKNA 179  
 Db 120 khrcppgfvvgagtpertntvckrcpdgffsnetskapcrkhtncsvfgllltqkna 179  
 Qy 180 THDNVCSGNREATOKCGIDVTICEAFAFRFAVPTKIIPNWLVSVLVDSPGTVKNAESVER 239  
 Db 180 thdnicsgnsestqkcgidvticeaaffrfaftpwnlsvldnlpgtkvnaesver 239  
 Qy 240 IKRRHSQEQTFQLLKLWKNQRDQEMVKYKIIQDIDCESSVQRHLGHSNLTTEQLLALM 299  
 Db 240 ikrhssqeqtfqlklwkhnqkdqkvkiiqdidlccensvqrhghanitfeqlrslm 299  
 Qy 300 ESLPGKKISPEIERTRKTCSSQOLKLLSLWRKNGDQDTLGLMYALKHLSHFPK 359  
 Db 300 eslpkgkvgaediektikackpsdqllkllslwrkngdgtlkglmhalkhsktyhfpk 359  
 Qy 360 TVTHSLRKTMRFLHSFTMYRYQKLFLEMIGNOVQSVKISCL 401  
 Db 360 tvtsqlkktirflhsftmykyqlflemignqvsvkiscsl 401

RESULT 9  
 Y43400  
 ID Y43400 standard; Protein; 401 AA.  
 XX Y43400;  
 AC Y43400;  
 DT 28-JAN-2000 (first entry)  
 DE Osteoprotegerin protein sequence.  
 KW Osteoprotegerin; OPG; human; cardiovascular disease; occlusion;  
 KW calcification; blood vessel; atherosclerosis; medial calcific sclerosis;  
 KW Monckeberg's arteriosclerosis; therapy.  
 XX Homo sapiens.  
 XX WO9953942-A1.  
 XX 28-OCT-1999.  
 XX 21-APR-1999; 99WO-US08793.  
 XX 23-APR-1998; 98US-0064832.  
 XX (AMGE-) AMGEN INC.  
 PA

XX Simonet S, Sarosi I;  
 XX WPI: 2000-013182/01.  
 DR N-PSDB; 237254.  
 XX  
 PT Treating and preventing cardiovascular diseases, especially  
 PT atherosclerosis and Monckeberg's arteriosclerosis  
 XX  
 PS Claim 9; Page 37-39; 43pp; English.  
 XX  
 CC This sequence represents the human osteoprotegerin (OPG). The invention  
 CC relates to a method of treating or preventing cardiovascular disease by  
 CC administering OPG. The method can be used to treat and prevent  
 CC cardiovascular diseases associated with occlusion and calcification of  
 CC blood vessels, especially atherosclerosis or Monckeberg's  
 CC arteriosclerosis, i.e. medial calcific sclerosis. Using OPG to treat or  
 CC prevent cardiovascular diseases provides an alternative to invasive  
 CC treatments. OPG can be used as a single therapeutic for prevention and  
 CC treatment of both osteoporosis and cardiovascular diseases.  
 XX  
 SQ Sequence 401 AA;

Query Match 86.8%; Score 1892; DB 21; Length 401;  
 Best Local Similarity 85.6%; Pred. No. 1e-154;  
 Matches 344; Conservative 26; Mismatches 30; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLLDI-IEWTTQETLPPKYLHYDPETHGQLLCKCAPGYLYKHQCHVRRK 59  
 DB 1 mnkllcca-lvfidisikwtgtetfpkylhydeetshqllcdkcpptgylkqhctakwk 59  
 QY 60 TLVCPCHSYTDSWHTSDCYVCSVKELQSVKQECNTHNRCVCEGRVLEIEFCL 119  
 DB 60 tvcapcdhyttdswhtsdclcyvcsvkelygkqecnrthnrvceckegrylelefc 119  
 QY 120 KHRSCPPGGVQAGTPERTVCKKCPDGFSGTSSKAPCIKHTNCSTFGLLLIQKNA 179  
 DB 120 khrcppgfvvqagtpertvckrcpdgffsnetsskapcrkhtncsvfgllltqkna 179  
 QY 180 THNVCSGNREATQKCGIDVTLCCEAFRFAVPTKIIPNWLVLDSLPOTKVNESVER 239  
 DB 180 thnicsgnsctqkcgidvtlceeafrfavptktpnwlslvldnlpotkvnaesver 239  
 QY 240 IKRRHSSQEQTFOLLKWLKHONRQEMWKIIQDIDLCSSVORHLGHSNLTTEQLALM 299  
 DB 240 ikrqhsseqtfqllkwlkhonrqemwkiiqdidlcensvgrhghnaltfeqlrslm 299  
 QY 300 ESLPGKKISPEETRTKTKSSQQLKLSLRIRKNGDQDTLGLMYALKHLKTSHPFK 359  
 DB 300 eslpgkkvgaediektikackpsdqllkllslwrknqgdqtklglmhalkhsctyhfpk 359  
 QY 360 TVTHSLRKTMRFLHSFTMYLYOKLFLEMIGNQVQSKISCL 401  
 DB 360 tvtqslkktirflhsftmylyoklflemignqdvqsvkisc 401

RESULT 10  
 W57635  
 ID W57635 standard; Protein; 401 AA.  
 XX  
 AC W57635;  
 XX  
 DT 27-AUG-1998 (first entry)  
 XX  
 DE Trl receptor protein.  
 XX  
 KW Trl receptor; tumour necrosis factor receptor-1; TNF-alpha; TNF-beta;  
 KW inhibitor; tumour growth; tumour necrosis; microorganism infection;  
 KW cellular differentiation stimulation; ionising radiation; septic shock;  
 KW anti-viral response; growth regulator; immune response; meningococemia;  
 KW autoimmune disease; apoptosis; inflammation; cachexia; cerebral malaria;  
 KW AIDS; therapy.

XX Homo sapiens.  
 XX WO9812344-A1.  
 XX  
 PD 26-MAR-1998.  
 XX  
 PF 18-SEP-1996; 96WO-US15003.  
 XX  
 PR 18-SEP-1996; 96WO-US15003.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Fleischmann RD, Greene JM, Ni J;  
 XX  
 DR WPI: 1998-217278/19.  
 DR N-PSDB; V24486.  
 XX  
 PT New isolated tumour necrosis factor receptor - useful for developing  
 PT products for treating, e.g. tumours, auto-immune disease(s), graft  
 PT rejection, apoptosis or inflammation  
 XX  
 PS Claim 13; Fig 1; 11pp; English.

CC This sequence is a tumour necrosis factor receptor-1 (Trl) receptor  
 CC of the invention. The DNA can be used to produce a recombinant host  
 CC cell by inserting it into a vector, which is then used to transfect the  
 CC host cell. The Trl receptor can bind both TNF-alpha and TNF-beta. Trl  
 CC receptor agonists can be used for inhibition of tumour growth and  
 CC necrosis of tumours. They can also be used to stimulate cellular  
 CC differentiation, e.g. T cell, fibroblasts or haematopoietic cell  
 CC differentiation, may be used to augment Trl's role in the host's defence  
 CC against microorganisms and prevent related disease. The agonists may  
 CC also be used to protect against the deleterious effects of ionising  
 CC radiation produced during a course of radiotherapy, e.g. denaturation of  
 CC enzymes, lipid peroxidation or DNA damage. The agonists may further be  
 CC used to mediate an anti-viral response, to regulate growth, to mediate  
 CC the immune response and to treat immunodeficiencies related to diseases  
 CC such as HIV. Antagonists to the Trl receptor may be used to treat  
 CC autoimmune diseases, e.g. graft versus host rejection and allograft  
 CC rejection, and T cell mediated autoimmune diseases. They may also be used  
 CC to prevent apoptosis. They may also be used to prevent cytotoxicity and  
 CC to treat septic shock, meningococemia, inflammation, bacterial  
 CC infections, cachexia, cerebral malaria or AIDS. The products can also be  
 CC used for diagnosing the above diseases.

Query Match 86.8%; Score 1891; DB 19; Length 401;  
 Best Local Similarity 85.3%; Pred. No. 1.2e-154;  
 Matches 343; Conservative 27; Mismatches 30; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLLDI-IEWTTQETLPPKYLHYDPETHGQLLCKCAPGYLYKHQCHVRRK 59  
 DB 1 mnkllcca-lvfidisikwtgtetfpkylhydeetshqllcdkcpptgylkqhctakwk 59  
 QY 60 TLVCPCHSYTDSWHTSDCYVCSVKELQSVKQECNTHNRCVCEGRVLEIEFCL 119  
 DB 60 tvcapcdhyttdswhtsdclcyvcsvkelygkqecnrthnrvceckegrylelefc 119  
 QY 120 KHRSCPPGGVQAGTPERTVCKKCPDGFSGTSSKAPCIKHTNCSTFGLLLIQKNA 179  
 DB 120 khrcppgfvvqagtpertvckrcpdgffsnetsskapcrkhtncsvfgllltqkna 179  
 QY 180 THNVCSGNREATQKCGIDVTLCCEAFRFAVPTKIIPNWLVLDSLPOTKVNESVER 239  
 DB 180 thenicsgnsctqkcgidvtlceeafrfavptktpnwlslvldnlpotkvnaesver 239  
 QY 240 IKRRHSSQEQTFOLLKWLKHONRQEMWKIIQDIDLCSSVORHLGHSNLTTEQLALM 299  
 DB 240 ikrqhsseqtfqllkwlkhonrqemwkiiqdidlcensvgrhghnaltfeqlrslm 299





XX This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CL95 in which the 19th Cys residue in the  
 CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 XX  
 SQ Sequence 401 AA;

Query Match 86.5%; Score 1885; DB 17; Length 401;  
 Best Local Similarity 85.3%; Pred. No. 4.1e-154;  
 Matches 343; Conservative 26; Mismatches 31; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLDI-IEWTTOETLPPKYLHYDPTGHQLLCKCAPGTYLKQHCTVRRK 59  
 DB 1 mnnllcca-lvldisikwtqetfppkylhydeetshqllcdkcpptylkqhtakwk 59

QY 60 TLCVPCPDHSDYSDWHTSDECVYSPVKELQSVKQECNTHNRVCEEGRYLEIEFCL 119  
 DB 60 tvcapcpdhytswhtsdeclcsyckelqyvkqecnrthnrvccekegryleiefcl 119

QY 120 KHRSCPPGSGVQAGTPERNVCKKCPDGFSGETSSKAPCIKHTNCSTFGLLLIQGNA 179  
 DB 120 khrrscppgsgvqagtpernvtckrcpdgffsnetsskapckhktncsvfllltqkna 179

QY 180 THDNVCSGNREATOKGIDVTLCBEAFRFAVPTKIIPNWLVSVDLSLPGTKVNAESVER 239  
 DB 180 thdnicsgsestqsgidvtlceaeaffravptkfpnlwslvdlpgtkvnaesver 239

QY 240 IKRRHSSQEQTFOLLKWKHQNDRQEMVKTIQDIDICESSVQRHLGHSNLTTEQLLALM 299  
 DB 240 ikrrhssqeqtfollkwwhqnkdqdivkikiqdidicensvqrhghantftqlslm 299

QY 300 ESLPGKKISPEIERTRKTKSSQQLKLLSLWRKNGDQDTLKGMLYALKHLKTSHPFK 359  
 DB 300 eslpqkkgvgaediektikackpsdqilkslwrlkngdgtlkgmlhalkhsktyhfpk 359

QY 360 TVTSLRKTMRFLHSFTMYRLYQKLFLEMIGNOVQSVKISCL 401  
 DB 360 tvtslktirflhsfmyklyqlflemignovqsvkiscl 401

RESULT 13  
 R99942  
 ID R99942 standard; Protein; 399 AA.  
 XX  
 AC R99942;  
 XX  
 XX 23-APR-1997 (first entry)  
 XX  
 XX Mutated OCIF, OCIF-CL.  
 XX  
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
 KW osteoporosis.  
 XX  
 XX Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /note= "Signal peptide"  
 FT Protein 22..399  
 FT /note= "Mature OCIF-CL"  
 XX  
 XX W09626217-A1.  
 XX  
 XX 29-AUG-1996.

XX 20-FEB-1996; 96WO-JP00374.  
 XX  
 XX 21-JUL-1995; 95JP-0207508.  
 PR 20-FEB-1995; 95JP-0054977.  
 XX  
 XX (SNOW ) SNOW BRAND MILK PROD CO LTD.  
 XX  
 XX Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
 XX  
 XX WPI; 1996-402320/40.  
 DR N-PSDB; T33172.  
 XX  
 XX DNA encoding osteoclastogenesis inhibitory factor protein - useful  
 PT for bone resorption control, esp. treatment of osteoporosis  
 PT  
 XX Claim 62; Page 117-119; 183pp; Japanese.  
 PS  
 XX This sequence represents a mutated version of the full length  
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
 CC sequence represents OCIF-CL in which amino acids 379-380 of the  
 CC mature OCIF protein are deleted. The OCIF of the invention  
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
 CC cation-exchangers or heparin and its activity is lowered after 10 mins  
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
 CC deg.C. OCIF is useful in the control of bone resorption and therefore  
 CC in the treatment and prevention of disorders of bone resorption, e.g.  
 CC osteoporosis.  
 XX  
 SQ Sequence 399 AA;

Query Match 86.4%; Score 1882; DB 17; Length 399;  
 Best Local Similarity 85.5%; Pred. No. 7.3e-154;  
 Matches 342; Conservative 26; Mismatches 30; Indels 2; Gaps 2;

QY 1 MNKWLCCALLVLDI-IEWTTOETLPPKYLHYDPTGHQLLCKCAPGTYLKQHCTVRRK 59  
 DB 1 mnnllcca-lvldisikwtqetfppkylhydeetshqllcdkcpptylkqhtakwk 59

QY 60 TLCVPCPDHSDYSDWHTSDECVYSPVKELQSVKQECNTHNRVCEEGRYLEIEFCL 119  
 DB 60 tvcapcpdhytswhtsdeclcsyckelqyvkqecnrthnrvccekegryleiefcl 119

QY 120 KHRSCPPGSGVQAGTPERNVCKKCPDGFSGETSSKAPCIKHTNCSTFGLLLIQGNA 179  
 DB 120 khrrscppgsgvqagtpernvtckrcpdgffsnetsskapckhktncsvfllltqkna 179

QY 180 THDNVCSGNREATOKGIDVTLCBEAFRFAVPTKIIPNWLVSVDLSLPGTKVNAESVER 239  
 DB 180 thdnicsgsestqsgidvtlceaeaffravptkfpnlwslvdlpgtkvnaesver 239

QY 240 IKRRHSSQEQTFOLLKWKHQNDRQEMVKTIQDIDICESSVQRHLGHSNLTTEQLLALM 299  
 DB 240 ikrrhssqeqtfollkwwhqnkdqdivkikiqdidicensvqrhghantftqlslm 299

QY 300 ESLPGKKISPEIERTRKTKSSQQLKLLSLWRKNGDQDTLKGMLYALKHLKTSHPFK 359  
 DB 300 eslpqkkgvgaediektikackpsdqilkslwrlkngdgtlkgmlhalkhsktyhfpk 359

QY 360 TVTSLRKTMRFLHSFTMYRLYQKLFLEMIGNOVQSVKIS 399  
 DB 360 tvtslktirflhsfmyklyqlflemignovqsvkis 399

RESULT 14  
 R99934  
 ID R99934 standard; Protein; 401 AA.  
 XX  
 XX R99934;  
 XX

DT 22-APR-1997 (first entry)  
XX Mutated OCIF, OCIF-C22S.  
DE Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis.  
XX Synthetic.  
OS Key Location/Qualifiers  
FH Peptide 1..21 /note= "Signal peptide"  
FT Protein 22..401  
FT /note= "Mature OCIF-C22S"  
FT Misc-difference 277  
FT /label= C22S  
XX WO9626217-A1.  
PN 29-AUG-1996.  
XX 20-FEB-1996; 96WO-JP00374.  
XX 21-JUL-1995; 95JP-0207508.  
XX 20-FEB-1995; 95JP-0054977.  
XX (SNOW ) SNOW BRAND MILK PROD CO LTD.  
XX Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
XX WPI; 1996-402320/40.  
XX N-PSDB; T33164.  
XX DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
XX Claim 38; Page 100-102; 183pp; Japanese.  
XX This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-C22S in which the 22nd Cys residue in the  
CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
XX osteoporosis.  
XX Sequence 401 AA;  
Query Match 86.3%; Score 1880; DB 17; Length 401;  
Best Local Similarity 85.1%; Pred. No. 1.1e-153;  
Matches 342; Conservative 26; Mismatches 32; Indels 2; Gaps 2;  
QY 1 MNKWLCCALLVLLDI-IEWTQETLPPKLYHYDPTGTHQLCDKCAPGTLYLKQCHTVRRK 59  
DB 1 mnnllccca-lvldisikwtgtfppkylhydeetshgllcdkcpptgylkghctakwk 59  
QY 60 TLCVPCPDHSTYDSWHTSECVCSFVCKELQSVQKCNRTNHRVCEEGRYLEIEFCL 119  
DB 60 tvacpdpdhytdswhtsdclcyevckelqykvqecnrtnhrvceckegryleiefcl 119  
QY 120 KHRSCPPGSGVQAGTPERNVCKKCPDGFPSGETSSKAPCIKHTNCSTFGLLLIQKNA 179  
DB 120 khrcpppgfvvqagtperntvckrcpdpffsnetsskpcrkhtncsvflllqkqna 179  
QY 180 THDNCVSGNREATQCGIDVLCIEAFRAFPVTKLIINWLSVLDLPGTKVNAESVER 239  
DB 180 thdnicsgnsstqcgidvltlceaeaffravptkfnwlsvldnlpgtkvnaesver 239

QY 240 IKRRHSSQETFQLLKLWKHQNRDOEMVKKIQQIDLCSSVQRHGLHSNLTTEOLLALM 299  
DB 240 ikrrhssqetfqlklwkhnkqkddivkikiqqidldlcensvqrhghlanltfeqlrslm 299  
QY 300 ESLPGKKISPEIEIRTKCKSSQELLKLLSLWRKNGDQDTLGLMYALKHLKTSHPFK 359  
DB 300 eslpgkkvgaediektikaskpsdqlnllslwrkngdtdlglmlhalkhsktyhfpk 359  
QY 360 TVTHSLRKTMRFLHSFTMYRLYQKLFLEMGVQSVKISCL 401  
DB 360 tvtqslkktirflhsftmyklyqkflflemignqvsvkisc1 401  
RESULT 15  
R99933  
ID R99933 standard; Protein; 401 AA.  
XX  
AC R99933;  
XX  
DT 22-APR-1997 (first entry)  
XX  
DE Mutated OCIF, OCIF-C21S.  
XX  
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;  
KW osteoporosis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..21 /note= "Signal peptide"  
FT Protein 22..401 /note= "Mature OCIF-C21S"  
FT Misc-difference 277 /label= C21S  
XX  
PN WO9626217-A1.  
XX  
PD 29-AUG-1996.  
XX  
PF 20-FEB-1996; 96WO-JP00374.  
XX  
PR 21-JUL-1995; 95JP-0207508.  
XX  
PR 20-FEB-1995; 95JP-0054977.  
XX  
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.  
XX  
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;  
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;  
XX  
DR WPI; 1996-402320/40.  
DR N-PSDB; T33163.  
XX  
XX DNA encoding osteoclastogenesis inhibitory factor protein - useful  
PT for bone resorption control, esp. treatment of osteoporosis  
XX  
XX Claim 35; Page 98-100; 183pp; Japanese.  
XX This sequence represents a mutated version of the full length  
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This  
CC sequence represents OCIF-C21S in which the 21st Cys residue in the  
CC mature OCIF protein is substituted by Ser. The OCIF of the invention  
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions  
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto  
CC cation-exchangers or heparin and its activity is lowered after 10 mins  
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90  
CC deg.C. OCIF is useful in the control of bone resorption and therefore  
CC in the treatment and prevention of disorders of bone resorption, e.g.  
XX osteoporosis.  
XX Sequence 401 AA;







DB 325 TISKTKGPKAPQVYTIPTPPKQMAKDKVSLTCTMTDFPDTITVWQWNGQPAENYKNT 384

QY 180 PVLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPCK 232

DB 385 QPIMTDGSIYFYSKLVKQSWNEAGNTFTCSVLHGLHNHTEKLSHSPCK 437

## RESULT 2

Q90544 PRELIMINARY; PRT; 684 AA.  
AC Q90544;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE NOVEL ANTIGEN RECEPTOR PRECURSOR  
OS Ginglymostoma cirratum (Nurse Shark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;  
OC Ginglymostomatidae; Ginglymostoma.  
OX NCBI\_TaxID=7801;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RX MEDLINE=95183140; PubMed=7877689;  
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,  
RA Flajnik M.F.;  
RT "A new antigen receptor gene family that undergoes rearrangement and  
RT extensive somatic diversification in sharks."  
RL Nature 374:168-173(1995).  
DR EMBL; U18701; AAB48195.1; -  
DR HSSP; P01857; 1FC1.  
DR INTERPRO; IPR003006; -  
DR PFAM; PF00047; ig; 6.  
DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_3.  
KW Signal.  
FT SIGNAL 1 18 POTENTIAL.  
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.  
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DEF CRC64;

Query Match 21.6%; Score 272.5; DB 13; Length 684;  
Best Local Similarity 33.3%; Pred. No. 9.4e-17;  
Matches 71; Conservative 35; Mismatches 98; Indels 9; Gaps 7;

QY 24 SVFLPPPKKDTLMISRPETVTCVVVDVSHEDPE-VKENWYVDGVEVHNATKPREQYN 82

DB 459 SVSLKLP-PFBEITWQQTATIVCEIV---YSDLENIKVFQVNGVGRKKGVETQNPWSG 514

QY 83 STYRVSVSLTVLHQLDNLNGEYKCVSNKALPAPIEKTISKAK-GQPREPQVYTLPPSKD 141

DB 515 SKSTIVSKLVYASWDSGTEVCLVDESELTPTPKASIRKANYSQMHPKPVLLHPSTD 574

QY 142 EL-TKNQVSLCLVKGFYPSDIWESNGO-PENNYKTTTPVLDSGDSFFLYSKLTVDK 199

DB 575 EIDTENSATLMCLATNFHFAEIIYQGWAMDTLLDSGYRTQVDSEKSGSFFVTLRLUTA 634

QY 200 SRWQGNVFCSCVMHEALHNHYTKSLSPCK 232

DB 635 AEWNSDITTSYCLVGHPSL-NRDLINKSTNKSNGK 666

## RESULT 3

Q9UP60 PRELIMINARY; PRT; 384 AA.  
AC Q9UP60;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE SNC73 PROTEIN.  
GN SNC73.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zheng S., Cao J., Cao W., Cai X., Geng L.;  
RT "Identification and characterization of SNC73, a gene which is down-  
RT regulated in colorectal cancer."  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF067420; AAC19365.1; -  
DR HSSP; P01825; 7FAB.  
DR INTERPRO; IPR003006; -  
DR PFAM; PF00047; ig; 3.  
DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;

Query Match 21.5%; Score 272; DB 4; Length 384;  
Best Local Similarity 31.2%; Pred. No. 5.1e-17;  
Matches 79; Conservative 33; Mismatches 109; Indels 32; Gaps 10;

QY 7 KTHTCP-----PCPAPELLGSPSVLEFPKPK-----DTLMISKTPEVT 45

DB 119 KHYTNPSQDVTVPCEVPTPTTSPST-PPTFSPSCCHPRLSLHRLPALEDLILLSEANLT 177

QY 46 CVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVSVSLTVLHODWLNKKEYK 105

DB 178 CTLGL-RDASGVFTTWPSSGK--SAVQGPPELDLCYSVSVLPGCAEPWNGKTF 234

QY 106 KCVSNKALPAPIEKTISKAKGPREPQVYTLPPSDELTKNO-VSLTCLVKGFYPSDIW 164

DB 235 CTAAYPESTKPLTATLSKS-GNTERPEVHLLPPPEELALNELVTLTCLARGFSPKDLV 293

QY 165 EWESNGO--PENNYKTTTPVLVD--SEFLYSLKLTVDKSRWQGNVFCSCVMHEALHN 219

DB 294 RWLQSGELPREKYLTVASROEPSQGTTFVAVTSLRVAEDWKKGDTFSCWVGHEALPL 353

QY 220 HYTKSLSPCK 232

DB 354 AFTQKTIDRLAGK 366

## RESULT 4

Q9NPP6 PRELIMINARY; PRT; 416 AA.  
AC Q9NPP6;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Auffray C., Ansoorge W., Ballabio A., Estivill X., Gibson K.,  
RA Lehrach H., Poustka A., Lundeberg J.;  
RT "The European IMAGE consortium for integrated Molecular analysis of  
RT human gene transcripts."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL389978; CAB97534.1; -  
FT NON\_TER 1  
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 21.2%; Score 268; DB 4; Length 416;  
Best Local Similarity 32.8%; Pred. No. 1.3e-16;  
Matches 77; Conservative 34; Mismatches 112; Indels 12; Gaps 9;

QY	4	SCDKTHTCPCPAPPELLGGSPVFLFPKPKDITLMISRTPEVTCVVDVSHEDPEVKFNMY	63
Db	170	QSDVTPCVPVPPPPPC -HPRLSLHRPALED-LLLGSEANLTCTLTGL-RDASGATFTWT	226
QY	64	VDGVEVHNATKPREBQNSTYRVVSVLTVLHODWLNKREYKCKYSNKALPAPIEKTISK	123
Db	227	PSSGK--SAVGQPPERDLGCVSVSVLPGCQAPWNHGETFTCTAAHPKLTPLTANITK	284
QY	124	AKGQPREPQYITLPPSRDELTKNQ-VSLTCLVKGYFSPDIAVWESNGO--PENNYKTPP	180
Db	285	S-GNTERPEVHLLPPPEELALNELVTLTCLARGFSPKDVLRWLQSGQELPREKYLTIWA	343
QY	181	PVLID-SDG--SFFLYSKLIVDKSRWQGNVFCFSVMHEALHNHYTKSLSLSPGK	232
Db	344	SROEPSQGTTFEAVTILRYAAEDWKKGDTFSCVMGHEALPLAFTQKTTIDRLAGK	398
RESULT	5		
Q9QX57			
ID	Q9QX57	PRELIMINARY;	PRT; 509 AA.
AC	Q9QX57;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)		
DE	BIT.		
GN	BIT.		
OS	Mus musculus (Mouse).		
OC	Eumariota; Metazoa; Chordata;		
OC	Kumalyota; Eutheria; Rodentia;		
OX	NCBI_TaxID=10090;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=129SV; TISSUE=LIVER;		
RC	Sano S., Ohnishi H., Kubota M.;		
RL	"Gene structure of mouse Bit/SHPS-1. ";		
RL	Biochem. J. 344:667-675(1999).		
DR	EMBL; AB024507; BAA89289.1; -		
DR	EMBL; AB024500; BAA89289.1; JOINED.		
DR	EMBL; AB024501; BAA89289.1; JOINED.		
DR	EMBL; AB024502; BAA89289.1; JOINED.		
DR	EMBL; AB024503; BAA89289.1; JOINED.		
DR	EMBL; AB024504; BAA89289.1; JOINED.		
DR	EMBL; AB024505; BAA89289.1; JOINED.		
DR	EMBL; AB024506; BAA89289.1; JOINED.		
DR	INTERPRO: IPR003006; -		
DR	PFAM; PF00047; ig; 3.		
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_1.		
SQL	SEQUENCE 509 AA; 55997 MW; 55023F2FCF1AA002B5 CRC64;		

Query Match	14.5%	Score 183;	DB 11;	Length 509;
Best Local Similarity	25.7%;	Pred. No.	1e-08;	
Matches	66;	Conservative	42;	Mismatches 107; Indels 42; Gaps 12;
QY	5	CDKTHTCPCCPAPELL--GGSEVLF-----PPK----	-PKDTLMISRTPEVTCCVVVDVSH--	53
Dd	121	CVKFKQGSEPDETIQSGGGEIVVLAKSPPEVSGPADRGIPDQKVNFCTC----	KSHGF	176
QY	54	EDPEVKENWYDVGEVEHNAAKTPREEQNSTYRVVSVLTVLHQDWLNCKEKVCVSNKAL	113	
Dd	177	SPRNITLKRFKDGOELHLEHTVNPSGNYSINSSITVRVY----LSMDEVHSKVICEVA	232	
QY	114	PAPTEKITSRAKGOPREPQQVTTLPSS-----RDELTKNOVSLTCLVGFGFSDIATVES	168	
Dd	233	HITLDR--SPLRGIANLSNIRVSFTVKVYQQSQSTSMNQMLTCAERFPEDQLIWLE	290	
QY	169	NGQPENNYKITPPYL--DSDGFFLYSKLTVDKRWWQQGNVFCFSVMHE----	ALHNHY-	222
Dd	291	NGNSVRN-DTPKNLTNTDGTYNWTSFLVNSSAHREDVVFTCOVKHDQOQPAITRHNV	348	
QY	223	-----QKSLSLSPG	231	
Dd	349	LGLAHSDDOGSMOTFPG	365	

RESULT 6

P97797 ID P97797 PRELIMINARY; PRT; 513 AA.  
AC P97797; P97796; Q35924; O88555; O88556;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE SUBSTRATE 1 PRECURSOR  
GN (P84) (SHP SUBSTRATE-1) (INHIBITORY RECEPTOR SHPS-1) (SHPS-1) (SHPS1).  
OS PTPN51 OR P84 OR SHPS-1 OR SHPS1 OR BIT.  
OC Mus musculus (Mouse).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10909;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=97223399; PubMed=9070220;  
RA Yamao T., Matozaki T., Amano K., Matsuda Y., Takahashi N., Ochi F.,  
RA Fujioaka Y., Kasuga M.,  
RT "Mouse and human SHPS-1: molecular cloning of cDNAs and chromosomal  
RT localization of genes";  
RL Biochem. Biophys. Res. Commun. 231:61-67(1997).  
RN [2]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-53: 422-424 AND 429-433.  
RC STRAIN=BALB/C; TISSUE=CEREBELLUM, AND BRAIN;  
CC Comu S., Weng W., Olinsky S., Ishwad P., Mi Z., Hempel J., Watkins S.,  
RA Lagenaur C.F., Narayanan V.;  
RT "The murine p84 neural adhesion molecule is SHPS-1, a member of the  
RT phosphatase-binding protein family";  
RL J. Neurosci. 17:8702-8710(1997).  
RN [3]  
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.  
RC STRAIN=C57BL/6; TISSUE=FETAL THYMUS;  
RX MEDLINE=98480500; PubMed=9712903;  
RA Veillette A., Thibaudaud E., Latour S.;  
RT "High expression of inhibitory receptor SHPS-1 and its association  
RT with protein tyrosine phosphatase SHP-1 in macrophages";  
RL J. Biol. Chem. 273:22719-22728(1998).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129SV; TISSUE=LIVER;  
RA Sano S., Ohnishi H., Kubota M.;  
RT "Gene structure of mouse BIT/SHPS-1.";  
RL Biochem. J. 344:667-675(1999).  
RN [5]  
RP FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.  
RX MEDLINE=90152134; PubMed=2303162;  
CC Chuang W., Lagenaur C.F.;  
RT "Central nervous system antigen p84 can serve as a substrate for  
RT neurite outgrowth";  
RL Dev. Biol. 137:219-232(1990).  
CC -1- FUNCTION: BINDS TO THE SH2 DOMAINS OF PROTEIN-TYROSINE PHOSPHATASE  
CC 2C (SHP-2) AFTER TYROSINE PHOSPHORYLATION INDUCED BY VARIOUS  
CC MITOGENS AND CELL ADHESION. PROTEIN SUPPORTS ADHESION OF  
CC CEREBELLAR NEURONS, NEURITE OUTGROWTH AND GLIAL CELL ATTACHMENT.  
CC MAY PLAY A KEY ROLE IN INTRACELLULAR SIGNALING DURING  
CC SYNAPTONEGENESIS AND IN SYNAPTIC FUNCTION. MAY ACT AS A DOCKING  
CC PROTEIN AND INDUCE TRANSLLOCATION OF SHP-2 FROM THE CYTOSOL TO THE  
CC PLASMA MEMBRANE. TYROSINE PHOSPHORYLATED PTPN51 FROM MACROPHAGES  
CC PRIMARILY ASSOCIATES WITH SHP-1.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- ALTERNATIVE PRODUCTS: THREE FORMS; ISOFORMS 1 (SHOWN HERE), 2 AND  
CC 3. SMALL, ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE CEREBRAL CORTEX, CEREBELLUM,  
CC SPINAL CORD, SPLEEN AND MACROPHAGES AND AT MUCH LOWER LEVEL IN THE  
CC KIDNEY, HEART, LIVER AND THYMUS. WITHIN THE CEREBELLUM, EXPRESSION  
CC IS DETECTED IN THE SYNAPTIC GLUMERULI, GRANULE CELL BODIES, GOLGI  
CC EPITHELIAL CELLS, OLFACTORY GLOBULE, NEURONS IN THE HIPPOCAMPUS AND  
CC THE DENTATE GYRUS. IN THE EYE, EXPRESSION IS FIRST SEEN IN THE  
CC INNER PLEXIFORM AND OPTIC FIBER LAYERS. LATER IN DEVELOPMENT

CC EXPRESSION IS ALSO SEEN IN THE OUTER PLEXIFORM, RETINA AND OUTER  
CC SEGMENTS OF THE PHOTORECEPTOR LAYER.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE EMBRYO FROM DAY 7. WHEN  
CC EXPRESSION IS RESTRICTED TO THE FLOOR PLATE REGION OF THE VENTRAL  
CC NEURAXIS. WIDESPREAD EXPRESSION IN A VARIETY OF CNS REGIONS BEGINS  
CC 2-5 DAYS AFTER BIRTH.

CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC TWO C1-LIKE AND ONE V-LIKE DOMAINS.

CC EMBL: D87968; BAA13521.1; -

CC EMBL: D87967; BAA13520.1; -

CC EMBL: U89694; AAB92591.1; -

CC EMBL: AF072543; AAC24886.1; -

CC EMBL: AF072544; AAC24887.1; -

CC EMBL: AB024507; BAA89290.1; -

CC EMBL: AB024500; BAA89290.1; JOINED.

CC EMBL: AB024501; BAA89290.1; JOINED.

CC EMBL: AB024502; BAA89290.1; JOINED.

CC EMBL: AB024503; BAA89290.1; JOINED.

CC EMBL: AB024504; BAA89290.1; JOINED.

CC EMBL: AB024505; BAA89290.1; JOINED.

CC EMBL: AB024506; BAA89290.1; JOINED.

CC MGD: MGI:108563; Ptpnsl.

CC INTERPRO: IPR003006; -

CC PFAM: PF00047; Ig; 3.

CC PROSITE: PS00290; IG\_MHC; UNKNOWN\_1.

CC Signal; Transmembrane; Alternative splicing; Immunoglobulin domain;

CC Glycoprotein; SH3-binding; Phosphorylation.

FT SIGNAL 1 31

FT CHAIN 32 513

FT PROTEIN TYROSINE PHOSPHATASE, NON-

FT RECEPTOR TYPE SUBSTRATE 1.

FT EXTRACELLULAR (POTENTIAL).

FT CYTOPLASMIC (POTENTIAL).

FT IG-LIKE V-TYPE DOMAIN.

FT IG-LIKE C1-TYPE DOMAIN.

FT IG-LIKE C1-TYPE DOMAIN.

FT SH2-BINDING (POTENTIAL).

FT SH2-BINDING (POTENTIAL).

FT SH2-BINDING (POTENTIAL).

FT SH2-BINDING (POTENTIAL).

FT SH2-BINDING (POTENTIAL).

FT SH2-BINDING (POTENTIAL).

FT PHOSPHORYLATION (BY CAMP) (POTENTIAL).

FT PHOSPHORYLATION (BY TYR-KINASES)

FT (POTENTIAL).

FT PHOSPHORYLATION (BY TYR-KINASES)

FT (POTENTIAL).

FT PHOSPHORYLATION (BY TYR-KINASES)

FT (POTENTIAL).

FT PHOSPHORYLATION (BY TYR-KINASES)

FT (POTENTIAL).

FT MOD\_RES 54

FT CARBOHYD 92

FT CARBOHYD 168

FT CARBOHYD 180

FT CARBOHYD 205

FT CARBOHYD 209

FT CARBOHYD 246

FT CARBOHYD 271

FT CARBOHYD 293

FT CARBOHYD 302

FT CARBOHYD 312

FT CARBOHYD 320

FT CARBOHYD 345

FT CARBOHYD 367

FT CARBOHYD 367

FT VARSPLIC 147

FT VARSPLIC 425

FT CONFLICT 10

FT CONFLICT 29

FT CONFLICT 67

FT CONFLICT 74

FT CONFLICT 83

FT CONFLICT 86

FT CONFLICT 90

FT CONFLICT 90

FT CONFLICT 91

FT CONFLICT 96

FT CONFLICT 114

FT CONFLICT 118

FT CONFLICT 126

FT CONFLICT 128

FT CONFLICT 128

FT CONFLICT 194

FT CONFLICT 224

FT CONFLICT 351

FT CONFLICT 365

FT CONFLICT 490

SQ SEQUENCE 513 AA; 56424 MW; 2AFDDDD5FA6C1EC7C CRC64;

Query Match 14.5%; Score 183; DB 11; Length 513;  
Best Local Similarity 25.7%; Pred. No. le-08;  
Matches 66; Conservative 42; Mismatches 107; Indels 42; Gaps 12;

Qy 5 CDK7HTCCPCAPPELL--GGPSVFLF-----PPK---PKDTLMISRTPEVTCVVDVSH-- 53

Db 121 CVKQKQSGPEPTEIQSGGTEVVVLAKPSPPEVSGPADRGIPDKVNFCTC-----KSHGF 176

Qy 54 EDPEVKFNWYDGVVEVHNKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKAL 113

Db 177 SPRNITLKWFQDGOELHLETTVPNSGKNVSYNISSTVRV---LNSMDVHVKVCEVA 232

Qy 114 PAPTEKTSKAKGPQREPQVVTLPSS-----RDELTKNOVSILCLVKGFPSPDIWVES 168

Db 233 HITLDR--SPLRGIANLSNFIRVSTVKVTCQSPSTSMQVNLCTRAERFYEPEDQLQILWE 290

Qy 169 NGQPNENYKTTTPVVL--DSGGSFFLYSKLTVDKSRWQGNVFCGVSVMHE---ALNHVT- 222

Db 291 NGNVSRL--DTPKNLTKNVTDGTYNTSLFLVNSSAHDVVFTCVQKHQQPAITRNHTV 348

Qy 223 -----OKSLSLSPG 231

Db 349 LGLAHSDDQSGSMQTFFP 365

RESULT 7

ID Q90524 PRELIMINARY; PRT; 268 AA.

AC Q90524;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE NOVEL ANTIGEN RECEPTOR (FRAGMENT).

OS Ginglymostoma cirratum (Nurse shark).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;

OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;

OC Ginglymostomatidae; Ginglymostoma.

OX NCBI\_TaxID=7801;

RN [1]

RP SEQUENCE OF 1-240 FROM N.A.

RC TISSUE=SPLEEN;

RX MEDLINE=95183140; PubMed=7877689;

RA Greenberg A.S., Avila D., Hughes M., Hughes E.C.,

RA Flajnik M.F.;

RT "A new antigen receptor gene family that undergoes rearrangement and

RT extensive somatic diversification in sharks.";

RL Nature 374:168-173(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=SPLEEN;

RA Greenberg A.S.;

RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

DR EMBL; U18680; AAB48352.1; -

DR INTERPRO: IPR003006; -

DR PFAM: PF00047; Ig; 2.

DR PROSITE: PS00290; IG\_MHC; UNKNOWN\_1.

FT NON\_TER 268

FT SEQUENCE 268 AA; 29579 MW; 0CDE524DD6E9FDC4 CRC64;



Query Match 14.3%; Score 181; DB 13; Length 268;  
Best Local Similarity 25.1%; Pred. No. 7e-09;  
Matches 62; Conservative 35; Mismatches 90; Indels 60; Gaps 9;

QY 23 PSVFLPPKPKDTLMISRTPE-----VTCVVVDVSHEDPEVKFNWYVDG-----66  
DB 14 PYVFD-----TARVDQTPQETTKETGESLSINCVRDSCALPSTYWNKKSGSTNEE 65  
QY 67 -----VEVNAKTKPREEQNSTYRVVSVLTVLHDLWLNCKEYKCKV--SNKALPAP 116  
DB 66 TTSKGRVYVETVNSGSKSFLRIND-----LTV-----EDSGTYRCVKYRKNNAWYDCG 113  
QY 117 IEKT-----ISKAKGQPREPQVYTL--PPSRDELTKNOVSLTCLVKGFPSPDIA 163  
DB 114 LEELDWIYVGGGTGYVNGPIPLSPPIVSLLSHSAEEORANGFVOLVCLISGYYPENIA 173  
QY 164 VEWESNGQP-ENNYKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYT 222  
DB 174 VSWQNTKTIITSGFATTSPVKTSNDFSCASLLKVLQEWMSRGSVYSCQVSHSATSSNQ 233  
QY 223 QKSLSLS 229  
DB 234 KEIRSTS 240

RESULT 8  
Q46632 PRELIMINARY; PRT; 506 AA.  
AC O46632;  
DT 01-JUN-1998 (TREMBlrel.. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE MID-1 ANTIGEN PRECURSOR.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98143722; PubMed=9485180;  
RA Brooke G.P., Parsons K.R., Howard C.J.;  
RT "Cloning of two members of the SIRP alpha family of protein tyrosine  
phosphatase binding proteins in cattle that are expressed on monocytes  
and a subpopulation of dendritic cells and which mediate binding to  
CD4 T cells";  
RL Eur. J. Immunol. 28:1-11(1998).  
DR EMBL; Y11045; CAA71942.1;  
DR INTERPRO; IPR003006;  
DR PFAM; PF00047; ig; 3.  
KW Signal.  
FT SIGNAL.  
FT CHAIN.  
FT CHAIN.  
SQ SEQUENCE 506 AA; 55117 MW; ACDDEA9C9708A82 CRC64;

Query Match 14.2%; Score 179; DB 6; Length 506;  
Best Local Similarity 27.0%; Pred. No. 2.3e-08;  
Matches 62; Conservative 36; Mismatches 108; Indels 24; Gaps 6;

QY 1 EPKSCDKTH--TCPPCPAPELGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSH--EDPE 57  
DB 133 EFKSGCGTHLTUNAKSPVPPVSGPTV-----RATPEQVNFCTSHGSPRN 179  
QY 58 VFNWYVDGVGEVNAKTKPREEQNSTYRVVSVLTVLHDLWLNCKEYKCKVSNKALPA-- 115  
DB 180 ISLWFKNGKNELASQTSVDPEDDNDYSINSTTKVLLATGDVHSQVCEVAHVTLGGP 239  
QY 116 PIEKTSKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPENN 175  
DB 240 PLRGANLSETIRVPP---TLEITRSPSAGNQNVTCQNKVFPRHLQLTLWLENGNMSRT 296

QY 176 YKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHE---ALHNHYT 222  
DB 297 EAASVLVENKDGTFNQTSWLLVNSSAHRVAVLTQCVHEDRQPAVSKNHT 346

RESULT 9  
O08907 PRELIMINARY; PRT; 509 AA.  
ID O08907;  
AC O08907;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE BRAIN IMMUNOLOGICAL-LIKE PROTEIN (BIT).  
GN BIT.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALE/C; TISSUE=BRAIN;  
RX MEDLINE=97230468; PubMed=9073522;  
RA Ohnishi H., Kubota M., Sano S.;  
RT "BIT (Bit) maps to mouse chromosome 2.";  
RL Genomics 40:504-506(1997).  
DR EMBL; D85785; BAA20376.1;  
DR MGD; MGI:107947; Bit.  
DR INTERPRO; IPR003006;  
DR PFAM; PF00047; ig; 3.  
DR PROSITE; PS00290; IG\_MHC; UNKNOWN.1.  
SQ SEQUENCE 509 AA; 56033 MW; 1E377ACE31BA13E CRC64;

Query Match 14.0%; Score 177; DB 11; Length 509;  
Best Local Similarity 27.7%; Pred. No. 3.6e-08;  
Matches 66; Conservative 38; Mismatches 96; Indels 38; Gaps 12;

QY 21 GGSVFLF---PPK---PKDTLMISRTPEVTCVVVDVSH--EDPEVKFNWYVDGVEVHN 71  
DB 139 GGEVTVLAKSPPEVSGPADRGIPDKVNFTC-----KSHGFSRNTLTKWFKDQGLHP 194  
QY 72 AKTKPREEQNSTYRVVS-VLTVLHDLWLNCKEYKCKVSNKALPAPIEKTISKAKGQPR- 129  
DB 195 LETTVNPSGKNVSYNISTTVRVVSLNMDVNSKVI-CEVAHITLDRSLRGIANLSNFRV 253  
QY 130 --EPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQPENNYKTPPVL--DS 185  
DB 254 SPTVKVTQOPPT---SMNQVNLTCRAERFPEDLQLIWLLENGNVSRN--DTPKNLTNT 307  
QY 186 DGSFFLYSKLTVDKSRWQGNVFCSCVMHE---ALHNHYT-----OKSLSLSPG 231  
DB 308 DGIYNTSLFLVNSSAHRVDVFTCVKHQDQQAIRNHTVLGLAHSSDQSGMQTFPG 365

RESULT 10  
O46631 PRELIMINARY; PRT; 506 AA.  
ID O46631;  
AC O46631;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE MYD-1 ANTIGEN PRECURSOR.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PERIPHERAL BLOOD;  
RX MEDLINE=98143722; PubMed=9485180;  
RA Brooke G.P., Parsons K.R., Howard C.J.;  
RT "Cloning of two members of the SIRP alpha family of protein tyrosine



DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE MHC CLASS II ALPHA SUBUNIT  
 GN MHC-AUHA-DBA  
 OS Aulonocara hansbaenschi  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthopterygii; Acanthopterygii; Perciformes; Labroidel;  
 OC Cichlidae; Aulonocara  
 OX NCBI\_TaxID=27781;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Murray B.W., Shintani S., Suelmann H., Klein J.;  
 RT "Major histocompatibility complex class II A genes in cichlid fishes:  
 RT identification, expression, linkage relationships, and haplotype  
 RT variation";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF212849; AAF5676.1; -;  
 KW MHC.  
 SQ SEQUENCE 238 AA; 26005 MW; 3D18BF05BF0A9C53 CRC64;

Query Match 13.4%; Score 169; DB 7; Length 238;  
 Best Local Similarity 29.9%; Pred. No. 7.6e-08;  
 Matches 63; Conservative 27; Mismatches 83; Indels 38; Gaps 10;  
 QY 44 VTCVVVDVSHEDPEV-----KENWYVDGVEVHNATKPREEQYNSTYRVVSVLTVLH 95  
 DB 11 VLVCSADVHLHDIQVGCSDSGEMYGIDGEEKAYADNKKEEITPQPPFVNPFP--Y 68  
 QY 96 QDWLNGKEYCKVSKNKPAPLPIETISKA-KGQPREP-----QVYTLPPSRDEL---TK 145  
 DB 69 QE-----GAYENAVANQICRENKRVGKMGKDYPLEQVAPSGVMYIT----RDEVEFEK 120  
 QY 146 NOVSLTCLVKGYPSPDIAYEVESNGQPPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQOG 205  
 DB 121 N--TLICHVTGTPAPVNVSWTKNGQKVTGSSINTPYPNKDGTFQTQISRLQFTP---QLG 175  
 QY 206 NVFSGVMEALHNHYTQ-----KSLSLSPG 231  
 DB 176 DIYSCAVQHALTQATKFEYVDSSAQSDPG 206

RESULT 14  
 QY0536 PRELIMINARY; PRT; 257 AA.  
 AC QY0536;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE NOVEL ANTIGEN RECEPTOR (FRAGMENT).  
 OS Ginglymostoma cirratum (Nurse shark).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;  
 OC Ginglymostomatidae; Ginglymostoma.  
 OX NCBI\_TaxID=7801;  
 RN [1]  
 RP SEQUENCE OF 1-242 FROM N.A.  
 RC TISSUE-SPLEEN;  
 RX MEDLINE=95183140; PubMed=7877689;  
 RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,  
 RA Flajnik M.F.;  
 RT "A new antigen receptor gene family that undergoes rearrangement and  
 RT extensive somatic diversification in sharks";  
 RL Nature 374:168-173(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SPLEEN;  
 RA Greenberg A.S.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U18693; AAB48364.1; -;  
 DR INTERPRO; IPR003006; -;  
 DR PFAM; PF00047; 19; 2.

DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
 FT NON\_TER 257  
 SQ SEQUENCE 257 AA; 28083 MW; BAD019A99D0BC491 CRC64;  
 Query Match 13.2%; Score 166.5; DB 13; Length 257;  
 Best Local Similarity 29.4%; Pred. No. 1.4e-07;  
 Matches 40; Conservative 21; Mismatches 64; Indels 11; Gaps 3;  
 QY 97 DWLNGKEYCKVSKNKPAPLPIETISKAQGPQPVYTL--PPSRDELTKNOVSLTCLV 154  
 DB 115 DWAVGGAYAAVGDTA-----VTVNFQPPSPPIVLLHSATEEQRANREVLVCLLI 166  
 QY 155 KGFYPSDIAEVESNGQP--ENNYKTPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSYM 213  
 DB 167 SGYPENIAVSWQKNTKTITSGFATISPVKTSNDFSCASLLKVPLOEWSRGVSVCQVS 226  
 QY 214 HEALHNHYTQKLSLS 229  
 DB 227 HSATSSNQKKEIRSTS 242

RESULT 15  
 QY0545 PRELIMINARY; PRT; 237 AA.  
 AC QY0545;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE NOVEL ANTIGEN RECEPTOR (FRAGMENT).  
 OS Ginglymostoma cirratum (Nurse shark).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;  
 OC Ginglymostomatidae; Ginglymostoma.  
 OX NCBI\_TaxID=7801;  
 RN [1]  
 RP SEQUENCE OF 1-219 FROM N.A.  
 RC TISSUE-SPLEEN;  
 RX MEDLINE=95183140; PubMed=7877689;  
 RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,  
 RA Flajnik M.F.;  
 RT "A new antigen receptor gene family that undergoes rearrangement and  
 RT extensive somatic diversification in sharks";  
 RL Nature 374:168-173(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-SPLEEN;  
 RA Greenberg A.S.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U18702; AAB48196.1; -;  
 DR INTERPRO; IPR003006; -;  
 DR PFAM; PF00047; 19; 2.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
 FT NON\_TER 237  
 SQ SEQUENCE 237 AA; 25831 MW; 9C002BD0B55C8656 CRC64;

Query Match 13.1%; Score 165; DB 13; Length 237;  
 Best Local Similarity 25.8%; Pred. No. 1.8e-07;  
 Matches 56; Conservative 32; Mismatches 83; Indels 46; Gaps 8;  
 QY 44 VTCVVVDV-----SHEDPEVKNWYVDGVEVHNATKPREEQYNSTYRV 87  
 DB 17 INCVLRDATSALGNTCWSRKKSLSTREETIAKGRGVETV---NSGSKSFLIND--- 69  
 QY 88 VSVLTVLHODWLNKEYCKVSKN-KALPAPIETISKA-----KGQPREQVYV 135  
 DB 70 ---LTV-----EDGTYRCGVGVYRCPCFLCSHLISYAGCGDGTVVTVNPGIPSPV 121  
 QY 136 L--PPSRDELTKNOVSLTCLVKGYPSPDIAYEVESNGQP--ENNYKTPPVLDSDGSFFLY 192  
 DB 122 LLHSATEEQRANRFVQLVCLISGYYPENIAVSWQKNTKTITSGFATISPVKTSNDFSCA 181

Qy 193 SKLTVDKSRWQGNVFCVMHEALHNYTQKSLSL 229  
| | | | | : | : | | | : : | |  
Db 182 SLLKVPLEWGRGVSCQVSHSATSSNORKEIRSTS 218

Search completed: March 1, 2001, 09:20:02  
Job time: 400 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:13:22 ; Search time 40.97 Seconds  
(without alignments)  
182.871 Million cell updates/sec

Title: US-09-389-782A-1

Perfect score: 1263

Sequence: 1 EPKCDKTHTPCPAPPELL.....MHEALHNHYTKLSLSPGK 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1263	100.0	330	1 GCL_HUMAN	P01857 homo sapien
2	1186	91.5	290	1 GC3_HUMAN	P01860 homo sapien
3	1145	90.7	326	1 GC2_HUMAN	P01859 homo sapien
4	1135	89.9	327	1 GC4_HUMAN	P01861 homo sapien
5	921	72.9	323	1 GC_RABIT	P01870 oryctolagus
6	896	70.9	329	1 GC2_CAVPO	P01862 cavia porce
7	845.5	66.9	329	1 GC3_MOUSE	P22436 mus musculus
8	842	66.7	333	1 GCB_RAT	P20761 rattus norv
9	834.5	66.1	398	1 GC3M_MOUSE	P03987 mus musculus
10	822.5	65.1	324	1 GCL_MOUSE	P01864 mus musculus
11	822.5	65.1	326	1 GCL_RAT	P20759 rattus norv
12	817.5	64.7	393	1 GC1M_MOUSE	P01869 mus musculus
13	809.5	64.1	329	1 GCC_RAT	P20762 rattus norv
14	809	64.1	330	1 GCRA_MOUSE	P01863 mus musculus
15	805	63.7	335	1 GCAB_MOUSE	P01864 mus musculus
16	804	63.7	399	1 GCAM_MOUSE	P01865 mus musculus
17	785.5	62.2	322	1 GCA_RAT	P20760 rattus norv
18	779.5	61.7	336	1 GCB_MOUSE	P01866 mus musculus
19	774.5	61.3	405	1 GCBN_MOUSE	P01867 mus musculus
20	358	28.3	454	1 MUC_HUMAN	P01871 homo sapien
21	357	28.3	455	1 MUC_MOUSE	P01872 mus musculus
22	356	28.2	391	1 MUCB_HUMAN	P04220 homo sapien
23	354	28.0	421	1 EPC_MOUSE	P06336 mus musculus
24	353	27.9	429	1 EPC_RAT	P01855 rattus norv
25	352	27.9	428	1 EPC_HUMAN	P01854 homo sapien
26	347	27.5	476	1 MUCN_MOUSE	P01873 mus musculus
27	345	27.3	458	1 MUC_RABIT	P03988 oryctolagus
28	335	26.5	479	1 MUCM_RABIT	P04221 oryctolagus
29	332	26.3	450	1 MUC_CANFA	P01874 canis famil
30	332	26.3	454	1 MUC_MESAU	P06337 mesocricetu
31	329	26.0	457	1 MUC_SUNNO	P20768 suncus muri
32	297	23.5	438	1 HVC2_HETFR	P23085 heterodontu
33	295.5	23.4	299	1 ALC_RABIT	P01879 oryctolagus

#### RESULT 1

ID	GCL_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG GAMMA-1 CHAIN C REGION.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RL	Biochemistry 9:3161-3170(1970).			
RN	[3]			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RL	Biochemistry 9:3171-3181(1970).			
RN	[4]			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RN	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[6]			
RP	DISULFIDE BONDS.			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X. Intrachain disulfide bonds."			

P23081 heterodontu  
P01871 gallus gall  
P23088 heterodontu  
P23086 heterodontu  
P01876 homo sapien  
P23084 heterodontu  
P20758 gorilla gor  
P01877 homo sapien  
P23735 ictalurus p  
P01878 mus musculus  
P01843 mus musculus  
P20767 rattus norv

#### ALIGNMENTS

34 287.5 22.8 438 1 HVCS\_HETFR  
35 287.5 22.8 446 1 MUC\_CHICK  
36 278.5 22.1 461 1 HVCM\_HETFR  
37 276.5 21.9 393 1 HVCS\_HETFR  
38 272 21.5 353 1 HVCS\_HUMAN  
39 272 21.5 370 1 HVCI\_HETFR  
40 268.5 21.3 353 1 ALC1\_GORGO  
41 268.5 21.2 340 1 ALC2\_HUMAN  
42 248.5 19.7 481 1 MUCM\_ICTPU  
43 245.5 19.4 344 1 ALC\_MOUSE  
44 177.5 14.1 105 1 LACI\_MOUSE  
45 172 13.6 104 1 LAC2\_RAT

Biochemistry 9:3188-3196(1970).  
 [7]  
 RN DISULFIDE BONDS.  
 RX MEDLINE=77070267; PubMed=1002129;  
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 RT "Rule of antibody structure. The primary structure of a monoclonal  
 RT IgG1 immunoglobulin (myeloma protein Nle), I: Purification and  
 RT characterization of the protein, the L- and H-chains, the  
 RT cyanogen bromide cleavage products, and the disulfide bridges.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
 [8]  
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RP MEDLINE=81208100; PubMed=7236608;  
 RA Deisenhofer J.;  
 RT "Crystallographic refinement and atomic models of a human Fc fragment  
 RT and its complex with fragment B of protein A from Staphylococcus  
 RT aureus at 2.9- and 2.8-A resolution.";  
 RL Biochemistry 20:2361-2370(1981).  
 CC -1- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE  
 CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)  
 CC MARKER & THE GIM (NON-1) MARKERS.  
 CC -1- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC 35,116,198,269 & 272.  
 CC -1- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
 CC 268-272.  
 CC -1- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC RESIDUES 198,267&272.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; J00228; AAC82527.1; ALT\_INIT.  
 DR PIR; A02146; GHU.  
 DR PDB; 1FC1; 15-JUL-92.  
 DR PDB; 1FC2; 15-JUL-92.  
 DR MIM; 147100; -  
 DR INTERPRO; IPR000495; -  
 DR INTERPRO; IPR003006; -  
 DR PFAM; PF00047; 1g; 3.  
 DR PROSITE; PS00290; 1g\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98  
 FT HINGE. 1  
 FT CH1. 98  
 FT CH2. 110  
 FT CH3. 223  
 FT DOMAIN 224 330  
 FT DISULFID 27 83  
 FT DISULFID 103 103  
 FT DISULFID 109 109  
 FT DISULFID 112 112  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT CARBOHYD 180 180  
 FT VARIANT 97 97  
 FT VARIANT 239 239  
 FT VARIANT 241 241  
 FT MOD\_RES 330 330  
 FT STRAND 123 126  
 FT HELIX 130 134  
 FT TURN 136 137  
 FT STRAND 141 148  
 FT STRAND 158 162  
 FT TURN 163 164

FT STRAND 165 166  
 FT STRAND 175 178  
 FT STRAND 183 190  
 FT HELIX 193 197  
 FT TURN 198 199  
 FT STRAND 202 206  
 FT STRAND 215 219  
 FT STRAND 227 227  
 FT STRAND 230 234  
 FT HELIX 238 240  
 FT TURN 241 242  
 FT STRAND 245 256  
 FT STRAND 260 266  
 FT TURN 267 268  
 FT STRAND 269 270  
 FT STRAND 274 276  
 FT STRAND 280 281  
 FT TURN 283 284  
 FT STRAND 287 296  
 FT HELIX 297 301  
 FT TURN 302 303  
 FT STRAND 306 312  
 FT TURN 313 314  
 FT TURN 316 317  
 FT STRAND 320 324  
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
 Query Match 100.0%; Score 1263; DB 1; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-95;  
 Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPKSCDTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
 DB 99 EPKSCDTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 158  
 QY 61 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120  
 DB 159 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 218  
 QY 121 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 180  
 DB 219 ISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 278  
 QY 181 PVLSDSGSFYLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPCK 232  
 DB 279 PVLSDSGSFYLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPCK 330  
 RESULT 2  
 GC3\_HUMAN  
 ID GC3\_HUMAN STANDARD; PRT; 290 AA.  
 AC P01860;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).  
 OS IGHG3.  
 GN Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE (DISEASE PROTEIN WIS).  
 RX MEDLINE=81021548; PubMed=6774747;  
 RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;  
 RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:  
 RT gamma 3 heavy-chain disease protein Wis.";  
 RL Biochemistry 19:4304-4308(1980).  
 RN [2]  
 RP REVISIONS TO 12-97 OF PROTEIN WIS.  
 RX MEDLINE=77118561; PubMed=402363;  
 RA Michaelisen T.E., Frangione B., Franklin E.C.;  
 RT "Primary structure of the 'hinge' region of human IgG3. Probable



RX MEDLINE=80114419; PubMed=118920;  
 RA Hofmann T., Parr D.M.;  
 RT "A note of the amino acid sequence of residues 381-391 of human  
 RL immunoglobulin gamma chains";  
 RN Mol. Immunol. 16:923-925(1979).  
 [5]  
 RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
 RA Hofmann T., Parr D.M.;  
 RL Submitted (MAR-1980) to the PIR data bank.  
 [6]  
 RP SEQUENCE OF 1-121 (DOT).  
 RX MEDLINE=95255298; PubMed=7737190;  
 RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
 RT "Characterization of the two unique human anti-flavin monoclonal  
 RL immunoglobulins";  
 RN Eur. J. Biochem. 228:886-893(1995).  
 [7]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=72033500; PubMed=4940472;  
 RA Milstein C., Frangione B.;  
 RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";  
 RN Biochem. J. 121:217-225(1971).  
 [8]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=69064124; PubMed=5782707;  
 RA Frangione B., Milstein C., Pink J.R.L.;  
 RT "Structural studies of immunoglobulin G.";  
 RN Nature 221:145-148(1969).  
 [9]  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; V00554; CAA23814.1; -;  
 DR EMBL; V00554; CAA23815.1; -;  
 DR EMBL; V00554; CAA23816.1; -;  
 DR EMBL; V00554; CAA23817.1; -;  
 DR PIR; A02148; G2HU.  
 DR MIM; 147110; -;  
 DR INTERPRO; IPR000495; -;  
 DR INTERPRO; IPR003006; -;  
 DR PEAM; PF00047; Ig; 3;  
 DR PROSITE; PS00290; IG\_MHC; 2;  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 219 CH2.  
 FT DOMAIN 220 326 CH3.  
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 140 200  
 FT DISULFID 246 304  
 FT VARIANT 60 60  
 S -> A (IN MYELOMA PROTEINS TIL & ZIE).  
 /FTid=VAR\_003889.  
 AT OR NEAR THE COMPLEMENT-BINDING SITE.  
 REMOVED POST-TRANSLATIONALLY (PROBABLE).  
 SEQUENCE 326 AA; 35884 MW; 8310878C6876CF9C CRC64;  
 Query Match 90.7%; Score 1145; DB 1; Length 326;  
 Best Local Similarity 91.4%; Pred. No. 1.le-85;  
 Matches 212; Conservative 9; Mismatches 7; Indels 4; Gaps 2;  
 QY 1 EPKSCDKTHTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60

Db 99 ERKCCVE---CPPCPAPP-VAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 154  
 Qy 61 NWYDGVVHNAAKTKPREEQYNSTYRVVSVLTVLDHQLWLNKCKVSKNKAIPAPIKTK 120  
 Db 155 NWYDGVVHNAAKTKPREEQYNSTYRVVSVLTVLDHQLWLNKCKVSKNKAIPAPIKTK 214  
 Qy 121 ISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTP 180  
 Db 215 ISKTKGQPREPQVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTP 274  
 Qy 181 PVLSDSGFFLYSKLTVDKSRWQGNFVSCVMHEALNHHYTKQSLSLSPGK 232  
 Db 275 PVLSDSGFFLYSKLTVDKSRWQGNFVSCVMHEALNHHYTKQSLSLSPGK 326  
 RESULT 4  
 GC4\_HUMAN  
 ID GC4\_HUMAN STANDARD; PRT; 327 AA.  
 AC P01861;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-4 CHAIN C REGION.  
 GN IGHG4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83157104; PubMed=6299662;  
 RA Ellison J.W., Buxbaum J.N., Hood L.E.;  
 RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene";  
 RL DNA 1:11-18(1981).  
 RN [2]  
 RP SEQUENCE OF 1-30 AND 81-326.  
 RX MEDLINE=70207560; PubMed=4192699;  
 RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;  
 RT "Human immunoglobulin subclasses. Partial amino acid sequence of the  
 RL constant region of a gamma 4 chain";  
 RL Biochem. J. 117:33-47(1970).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; K01316; AAB59394.1; ALT\_INIT.  
 DR PIR; A02150; G4HU.  
 DR MIM; 147130; -;  
 DR INTERPRO; IPR000495; -;  
 DR INTERPRO; IPR003006; -;  
 DR PFAM; PF00047; Ig; 3;  
 DR PROSITE; PS00290; IG\_MHC; 2;  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 220 CH2.  
 FT DOMAIN 221 327 CH3.  
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 83  
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 141 201  
 FT DISULFID 247 305  
 SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;  
 Query Match 89.9%; Score 1135; DB 1; Length 327;



Best Local Similarity 93.7%; Pred. No. 7.1e-85;  
Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 11 CDPAPPELLGSPVFLFPKPKDTLMISRTPEVTVVVDVSHEDPEVFNWYVDGVEVH 70  
Db 106 CPSPAPPELLGSPVFLFPKPKDTLMISRTPEVTVVVDVSHEDPEVFNWYVDGVEVH 165

QY 71 NAKTPREEQNSTYRVSVTLVHODWLNKGYKCKVSNKALPAPIEKTISKAKQPRE 130  
Db 166 NAKTPREEQNSTYRVSVTLVHODWLNKGYKCKVSNKALPAPIEKTISKAKQPRE 225

QY 131 PQVYTLPPSRDELTKNQVSLTCLVKGFPYSDIAVENESNGQPENNYKTPPVLDSDGSEFF 190  
Db 226 PQVYTLPPSRDELTKNQVSLTCLVKGFPYSDIAVENESNGQPENNYKTPPVLDSDGSEFF 285

QY 191 LYSKLTVDKSRWQGNVSCSVNHEALHNHYTKLSLSPGK 232  
Db 286 LYSRLTVDKSRWQGNVSCSVNHEALHNHYTKLSLSPGK 327

RESULT 5  
GC\_RABIT STANDARD; PRT; 323 AA.  
ID GC\_RABIT  
AC P01870;  
DF 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA CHAIN C REGION.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84030930; PubMed=6313520;  
RA Bernstein K.E., Alexander C.B., Mage R.G.;  
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant  
F-1 haplotype."  
RL Immunogenetics 18:387-397(1983).  
RN [2]  
RP SEQUENCE OF 1-128.  
RX MEDLINE=76135469; PubMed=1243651;  
RA Pratt D.M., Mole L.E.;  
RT "Sequence studies on the constant region of the Fd sections of rabbit  
immunoglobulin G of different allotype."  
RL Biochem. J. 151:337-349(1975).  
RN [3]  
RP SEQUENCE OF 88-266 FROM N.A.  
RX MEDLINE=83299917; PubMed=6193512;  
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;  
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma  
heavy chain and identification of two genomic C gamma genes."  
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).  
RN [4]  
RP SEQUENCE OF 132-161.  
RX MEDLINE=70110015; PubMed=5461106;  
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;  
RT "Sequence studies of the Fd section of the heavy chain of rabbit  
immunoglobulin G."  
RL Biochem. J. 116:249-259(1970).  
RN [5]  
RP SEQUENCE OF 129-131 AND 155-322.  
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;  
RL (In) Killander J. (eds.);  
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,  
Stockholm (1967).  
CC -I- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,  
104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15  
MARKERS AND REF.5 THE E15 MARKER.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial  
entires requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to a license@isb-sib.ch).

CC EMBL; M16426; AAA31289.1; -.  
CC PIR; A02161; GHRB.  
CC INTERPRO; IPR000495; -.  
CC INTERPRO; IPR003006; -.  
CC PFAM; PF00047; ig; 3.  
CC PROSITE; PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1  
FT VARIANT 104 104 T -> M (IN D11 MARKER).  
FT VARIANT 185 185 T -> A (IN E15 MARKER).  
FT CONFLICT 48 48 V -> E (IN REF. 2).  
FT CONFLICT 71 71 V -> VPV (IN REF. 2).  
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).  
FT CONFLICT 173 173 N -> D (IN REF. 5).  
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).  
FT CONFLICT 201 201 N -> D (IN REF. 5).  
FT CONFLICT 218 218 Q -> E (IN REF. 5).  
FT CONFLICT 233 233 E -> Q (IN REF. 5).  
FT CONFLICT 246 246 N -> D (IN REF. 5).  
FT CONFLICT 256 256 E -> G (IN REF. 5).  
FT CONFLICT 260 260 N -> D (IN REF. 5).  
FT CONFLICT 266 266 N -> D (IN REF. 5).  
FT CONFLICT 280 280 Y -> W (IN REF. 5).  
FT CONFLICT 284 284 N -> S (IN REF. 5).  
SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 72.9%; Score 921; DB 1; Length 323;  
Best Local Similarity 67.3%; Pred. No. 1.4e-67;  
Matches 167; Conservative 31; Mismatches 34; Indels 16; Gaps 2;

QY 1 EPKSCDKT-----TC--PPCAPPELLGSPVFLFPKPKDTLMISRTPEV 44  
Db 76 QPVTCTNVAHPATNTKVDKTVAPSTCSKPTCPPELLGSPVFLFPKPKDTLMISRTPEV 135

QY 45 TCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREEQNSTYRVSVTLVHODWLNKGY 104  
Db 136 TCVVVDVSDQDPEVQFTWYINNEQVTRAPPLREQQFNSTIRVSTLPIHODWLNKGY 195

QY 105 KCVSNKALPAPIEKTISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAV 164  
Db 196 KCVSNKALPAPIEKTISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAV 255

QY 165 EWESNGQPNENNYKTPPVLDSDGSEFFLYSKLTVDKSRWQGNVSCSVNHEALHNHYTK 224  
Db 256 EWEKNGKAEDNYKTPPVLDSDGSEFFLYSKLTVDKSRWQGNVSCSVNHEALHNHYTK 315

QY 225 SLSLSPGK 232  
Db 316 SLSLSPGK 323

RESULT 6  
GC2\_CAVPO STANDARD; PRT; 329 AA.  
ID GC2\_CAVPO  
AC P01862;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2 CHAIN C REGION.  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
RN [1]  
RP SEQUENCE OF 1-3.  
RA Trischmann T.M.;  
RL Submitted (APR-1975) to the PIR data bank.  
RN [2]  
RP SEQUENCE OF 4-68.



DE IG GAMMA-2B CHAIN C REGION.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89232738; PubMed=3149946;  
 RA Brueggemann M.;  
 RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family";  
 RL Gene 74:473-482(1988).  
 DR PIR: PS0018; PS0018.  
 DR INTERPRO: IPR000495;  
 DR PFAM: PF00047; ig; 3.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1  
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 80  
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 147 207  
 FT DISULFID 253 311  
 SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 66.7%; Score 842; DB 1; Length 333;  
 Best Local Similarity 65.4%; Pred. No. 3.6e-61;  
 Matches 149; Conservative 33; Mismatches 46; Indels 0; Gaps 0;

QY 5 CDKTHPCPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENYVY 64  
 DB 106 CPTCTCHCKCPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENYVY 165  
 QY 65 DGEVHNKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISK 124  
 DB 166 NVEVHTAQTPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISK 225  
 QY 125 KGQPREQYVTLPPSDELTKNQSVLTCLVKGYFSPDAVWESNGOPENNYKTPPVLD 184  
 DB 226 KGLVRPQVYVMPPEQTEQVSLTCLTSGFLPDIGVWETSNGHIEKNYKNTPEVMD 285  
 QY 185 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYTKQSLSPGK 232  
 DB 286 SDGSFFMYSKLNVRSRWDSPAPFCVSVVHEGLHNHHVKSISRPPGK 333

RESULT 9  
 GC3M\_MOUSE STANDARD; PRT; 398 AA.  
 AC P03987;  
 DT 23-OCT-1986 (Rel. 02, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-3 CHAIN C REGION, MEMBRANE-BOUND FORM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85027161; PubMed=6092053;  
 RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,  
 RA Tucker P.W., Blattner F.R.;  
 RT "Structural analysis of the murine IgG3 constant region gene";  
 RL EMBO J. 3:2041-2046(1984).  
 RN [2]  
 RP SEQUENCE OF 328-398 FROM N.A.  
 RX MEDLINE=84041483; PubMed=6314238;  
 RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,  
 RA Wall R.;  
 RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene

RT segment.";  
 RL Nucleic Acids Res. 11:6775-6785(1983).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----

DR EMBL: J00451; AAB59655.1;  
 DR EMBL: V01526; AAB24767.1; ALT\_SEQ.  
 DR PIR: A02155; G3MSM.  
 DR INTERPRO: IPR000495;  
 DR INTERPRO: IPR003006;  
 DR PFAM: PF00047; ig; 3.  
 DR PROSITE: PS00290; IG\_MHC; 1.

KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 Transmembrane; Alternative splicing.  
 FT NON\_TER 1  
 FT DOMAIN 1 97 CHI.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 223 CH2.  
 FT DOMAIN 224 327 CH3.  
 FT TRANSMEM 346 362 POTENTIAL.  
 FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).  
 FT CONFLICT 333 333 E -> G (IN REF. 2).  
 FT CONFLICT 342 342 E -> Q (IN REF. 2).  
 FT CONFLICT 388 388 P -> F (IN REF. 2).  
 SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match 66.1%; Score 834.5; DB 1; Length 398;  
 Best Local Similarity 66.7%; Pred. No. 1.8e-60;  
 Matches 148; Conservative 34; Mismatches 37; Indels 3; Gaps 1;

QY 12 PP---CPAPELGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENYVYDGE 68  
 DB 106 PPGSSCPGPNLGGPSVFLFPKPKDMLSLPKVTCVVVDVSEDDPDVHVSFVNDKE 165  
 QY 69 VHNKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAKGP 128  
 DB 166 VHTAWTPQREAFNSIFRVVSALPIQHDWNGKEYCKVSNKALPAPIETISKAKGRA 225  
 QY 129 REPQVTLPPSRDELTKNQSVLTCLVKGYFSPDAVWESNGOPENNYKTPPVLDSDGS 188  
 DB 226 QTPQVYTIPTPPREQMSKKVSLTCLTNTFTSEATSEVWERNGELEQDYKNTPTLLSDGT 285  
 QY 189 FFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYTKQSLSP 230  
 DB 286 YFLYSKLTVDTSWLOGEIFTCSVWHEALHNHHTQKNLSRSP 327

RESULT 10  
 GC1\_MOUSE STANDARD; PRT; 324 AA.  
 AC P01868;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE IG GAMMA-1 CHAIN C REGION.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80045036; PubMed=115593;  
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
 RA Takahashi N., Mano Y.;  
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin  
 RT gamma 1 chain gene";  
 RL Cell 18:559-568(1979).

[2]  
 RN SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).  
 RX MEDLINE=80202559; PubMed=6769752;  
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,  
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;  
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences  
 RT cloned in a bacterial plasmid.";  
 RL Gene 9:87-97(1980).  
 RN [3]  
 RN SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=80012837; PubMed=113776;  
 RA Rogers J., Clarke P., Salsner W.;  
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin  
 RT heavy chain.";  
 RL Nucleic Acids Res. 6:3305-3321(1979).  
 RN [4]  
 RN SEQUENCE (MYELOMA PROTEIN MOPC 21).  
 RX MEDLINE=78242288; PubMed=98524;  
 RA Adetudbo K.;  
 RT "Evolution of immunoglobulin subclasses. Primary structure of a  
 RT murine myeloma gamma chain.";  
 RL J. Biol. Chem. 253:6068-6075(1978).  
 RN [5]  
 RN DISULFIDE BONDS (MOPC 21).  
 RX MEDLINE=73008889; PubMed=5073237;  
 RA Svasti J., Milstein C.;  
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";  
 RL Biochem. J. 126:837-850(1972).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; V00793; CAA24172.1; -;  
 DR EMBL; V00793; CAA24173.1; -;  
 DR EMBL; V00793; CAA24174.1; -;  
 DR EMBL; V00793; CAA24175.1; -;  
 DR EMBL; V00795; CAA24176.1; -;  
 DR PIR; A02159; GIMS.  
 DR MGD; MGI:96446; Igh-4.  
 DR INTERPRO; IPR000495; -;  
 DR INTERPRO; IPR003006; -;  
 DR PFAM; PF00047; Ig; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW Alternative splicing.  
 FT NON\_TER 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 110 HINGE.  
 FT DOMAIN 111 217 CH2.  
 FT DOMAIN 218 324 CH3.  
 FT DISULFID 27 82  
 FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 138 198  
 FT CARBOHYD 174 174  
 FT DISULFID 244 302  
 FT MOD\_RES 324 324  
 FT CONFLICT 276 276 N -> D (IN REF. 3).  
 FT CONFLICT 278 278 N -> D (IN REF. 3).  
 SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 65.1%; Score 822.5; DB 1; Length 324;  
 Best Local Similarity 60.9%; Pred. No. 1.3e-59;  
 Matches 142; Conservative 45; Mismatches 37; Indels 9; Gaps 3;

QY 2 PKCDKTHTCPP--CPAPELLGSPVFLFPKPKDMLTMSRPTVTCVVVDVSHEDPEVK 59  
 Db 99 PRDCG---CKPCICTVPEV---SSVFPPKPKDVLTLITLTPKVTCTVVVDISKDDPEVQ 157  
 QY 60 FNWYVDGVEVHNATKPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSKNKAIPAPIEK 119  
 Db 152 FSWFVDDVEVHTAQTQPREQFNSTFRSVSELPIMHQDWLNGKEYCKVSKNKAIPAPIEK 217  
 QY 120 TISKAGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 179  
 Db 212 TISTKGRKAPQVYITPPPKAEQMAKDKVSLTCLMTEFFPEDITVEWQNGQPAENYKNT 271  
 QY 180 PVLDSGDSFELYSLKTVDKSRWQGNVFCVMHEALHNHYTQKSLSLSPGK 232  
 Db 272 QPIMNTGSGFYVSKLVNPKVQKSNWEAGNTFTCSVLHLEGLHNHTKLSLSHPGK 324  
 RESULT 11  
 GCL\_RAT  
 ID GCL\_RAT STANDARD; PRT; 326 AA.  
 AC P20759;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DE 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-1 CHAIN C REGION.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=89232738; PubMed=3149946;  
 RX Brueggemann M.;  
 RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";  
 RL Gene 74:473-482(1988).  
 DR PIR; PS0017; PS0017.  
 DR INTERPRO; IPR000495; -;  
 DR INTERPRO; IPR003006; -;  
 DR PFAM; PF00047; Ig; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
 FT NON\_TER 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 112 HINGE.  
 FT DOMAIN 113 219 CH2.  
 FT DOMAIN 220 326 CH3.  
 FT DISULFID 27 82  
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 140 200  
 FT DISULFID 246 304  
 FT CARBOHYD 176 176  
 FT CARBOHYD 176 176 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 65.1%; Score 822.5; DB 1; Length 326;  
 Best Local Similarity 61.5%; Pred. No. 1.3e-59;  
 Matches 144; Conservative 43; Mismatches 38; Indels 9; Gaps 3;  
 QY 2 PKCDKTHTCPPAPPELLGG---PSVFLFPKPKDMLTMSRPTVTCVVVDVSHEDPEV 58  
 Db 99 PRNCG--GDCKPC----ICTGSEVSSVFIFPPKPKDVLTLITLTPKVTCTVVVDISQDPEV 152  
 QY 59 KFNWYVDGVEVHNATKPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSKNKAIPAPIEK 118  
 Db 153 HFSWFVDDVEVHTAQTQPREQFNSTFRSVSELPILHODWLNNGKEYCKVSKNKAIPAPIEK 212  
 QY 119 KTISKAGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 178  
 Db 213 KTISKPEGRTOVPHYTMSPTKEEMTQNEVITCMVKGFFYPDDIVVEWQNGQPAENYKNT 272  
 QY 179 TTPVLDSGDSFELYSLKTVDKSRWQGNVFCVMHEALHNHYTQKSLSLSPGK 232

FT	DOMAIN	1	97	CHI.
FT	DOMAIN	98	110	HINGE.
FT	DOMAIN	111	217	CH2.
FT	DOMAIN	218	324	CH3.
FT	DISULFID	27	82	
FT	DISULFID	102	102	INTERCHAIN (WITH A LIGHT CHAIN).
FT	DISULFID	104	104	INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID	107	107	INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID	109	109	INTERCHAIN (WITH A HEAVY CHAIN).
FT	DISULFID	138	198	
FT	CARBOHYD	174	174	N-LINKED (GLCNAC. . .).
FT	DISULFID	244	302	POTENTIAL.
FT	TRANSEM	340	357	CYTOPASMIC (POTENTIAL).
FT	DOMAIN	358	393	
FT	SEQUENCE	393 AA; 43386 MW; 4CC88343B7ALCE27 CRC64;		
Query Match		64.7%; Score 817.5; DB 1; Length 398;		
Best Local Similarity		60.8%; Pred. No. 4.2e-59;		
Matches 141; Conservative		45; Mismatches 37; Indels 9; Gaps 3;		
QY	2 PKSCDKTHTCPP--CPAPELLGGPSVFLFPKPKDGLTLMISRTPEVTCVVDVSHEDPEVK	59		
DB	99 PROCG-----CKPICITVPEV---SSVFIFFPKADVLTITLTPKVTGVVDISMDPEVQ	151		
QY	60 FNYIVDGVVEYHNKTKPREQYNTYRWVSVLTVLHODWLNKREYKCKVSNKALPAPIEK	119		
DB	152 FSNFVDDVEVHTAQTPREQFNSTFRSVSELPIMHQDLNKEFKCRVNSAAPAPIEK	211		
QY	120 TISKAGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVESNGQPNNYKTT	179		
DB	212 TISKTGRKPAQYVITPPPEKQAKDKVSLTCMITDFFPDIVQWNGQPAENYKNT	271		
QY	180 PPVLDSGSEFLYSKLVDRKSMQOGNVCSCVMHEALHNHYTKSLSPG	231		
DB	272 QPIMNTNGSYFYVSKLVNQKSNWEAGNTFTCSVLHEGLHNHHTKSLSHSPG	323		
RESULT 13				
GCC_RAT	STANDARD; PRT; 329 AA.			
AD	P20762;			
IC	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG GAMMA-2C CHAIN C REGION.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=88166903; PubMed=3127222;			
RA	Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;			
RT	"Sequence of a rat immunoglobulin gamma 2c heavy chain constant			
RT	region cDNA: extensive homology to mouse gamma 3.";			
RL	Eur. J. Immunol. 18:317-319(1988).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL; X07189; CAA30169.1; -			
DR	PIR; S00847; S00847.			
DR	INTERPRO; IPR000495; -			
DR	INTERPRO; IPR003006; -			
DR	PFAM; PF00047; Ig; 3.			
DR	PROSITE; PS00290; IG_MHC; 1.			
DR	Immunoglobulin domain; Immunoglobulin C region.			
NON TER	1			

immunoglobulin:amino-acid sequence of the Fc fragment. Implications for the evolution of immunoglobulin structure and function.";  
Eur. J. Biochem. 43:423-435(1974).  
[5]  
DISULFIDE BONDS.  
MEDLINE=73056887; PubMed=4565406;  
de Preval C., Fougereau M.;  
"Determination of the primary structure of a mouse gamma G2a immunoglobulin. Identification of the disulfide bridges.";  
Eur. J. Biochem. 30:452-462(1972).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; V00798; CAA24178.1; -  
PIR; A02152; G2MSA.  
INTERPRO; IPR000495; -  
INTERPRO; IPR003006; -  
PFAM; PF00047; Ig; 3.  
PROSITE; PS00290; IG\_MHC; 1.  
Immunoglobulin domain; Immunoglobulin C region.  
NON\_TER 1  
DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
DISULFID 27 82  
DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
DISULFID 144 204  
DISULFID 250 308  
MOD\_RES 330 330  
REMOVED POST-TRANSLATIONALLY.  
SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;  
  
Query Match 64.1%; Score 809; DB 1; Length 330;  
Best Local Similarity 65.2%; Pred. No. 1,7e-58;  
Matches 146; Conservative 30; Mismatches 46; Indels 2; Gaps 1;  
  
QY 11 CPP--CPAPELLGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEYKFNWYDGE 68  
Db 107 CPPCKCPAPNLLGGPSVFIPPKIKDVLMSLSPIVTCVVVDVSEDDPDVQISWFVNVE 166  
  
QY 69 VHNARTKPREQYNSTRYRVSVLTVLHODLNGCKEYKCKVSNKALPAPIEKTISKAKGQP 128  
Db 167 VTAQQTITREDYNSLTRVYSALPIQHQDMSGKEFKCKVNNKDLPAPIERTISKPGSV 226  
  
QY 129 REPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLDSDGS 188  
Db 227 RAPQVYVLPPPEEMTKQVLTLCWYDFWPEDIIYEWYNNGTSLNKTNPVLDSDGS 286  
  
QY 189 FFYLSKLTVDKRWQOQGNVFCSSVMHEALHNHYTKQSLSLSPGK 232  
Db 287 YFMYSKLRYEKNWVRNYSYSCSVHVEGLNHHHTKSFSTRTPCK 330  
  
RESULT 15  
GCAB\_MOUSE STANDARD; PRT; 335 AA.  
ID GCAB\_MOUSE  
AC P01864;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2A CHAIN C REGION, B ALLELE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
RN Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
RP [1]  
RP SEQUENCE FROM N.A.  
RP STRAIN=C57BL/6;

RX MEDLINE-82037861; PubMed-6170065;  
RA Schreier P.H., Bothwell A.L.M., Mueller-Hill B., Baltimore D.;  
RT "Multiple differences between the nucleic acid sequences of the  
RL Igg2aa and Igg2ab alleles of the mouse."  
RL Proc. Natl. Acad. Sci. U.S.A. 78:4495-4499(1981).  
RN [2]  
RP SEQUENCE.  
RX MEDLINE-82037777; PubMed-6794027;  
RA Dognin M.J., Lauwereys M., Strosberg A.D.;  
RT "Multiple amino acid substitutions between murine gamma 2a heavy  
chain Fc regions of Igla and Ig1b allotypic forms."  
RL Proc. Natl. Acad. Sci. U.S.A. 78:4031-4035(1981).  
CC -!- MISCELLANEOUS: THE SEQUENCE DIFFERS FROM THAT OF THE A ALLELE,  
CC FROM BALB/C MICE, AT 15% OF THE POSITIONS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: J00479; -: NOT\_ANNOTATED\_CDS.  
DR PIR: A02153; G2MSAB.  
DR INTERPRO: IPR000495; -.  
DR INTERPRO: IPR003006; -.  
DR PFAM: PF00047; ig; 3.  
DR PROSITE: PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1  
SQ SEQUENCE 335 AA; 36596 MW; FA3382792CBB13C6 CRC64;  
  
Query Match 63.7%; Score 805; DB 1; Length 335;  
Best Local Similarity 62.3%; Pred. No. 3.6e-58;  
Matches 142; Conservative 38; Mismatches 48; Indels 0; Gaps 0;  
  
QY 5 CDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 64  
DB 108 CPHQVRVPCAAPDLGGPSVFIFPPKIKDVLMLISLSPWVTCVVVDVSEDDPDVQISWV 167  
QY 65 DGEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTITSKA 124  
DB 168 NNVEVHTAQPTHTREDYNSLTRVVSALPIQHQMWSGKEFKCKVNNRALSPIEKTISK 227  
QY 125 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSDIAVWESNGQPENNYKTTPPVLD 184  
DB 228 RGPVRAFPQVYVLPPEAEEMTKKEFSLTCLMITGFLPAEIAVDWTSNGRTEQNKNTATVLD 287  
QY 185 SDGSFFLYSKLTVDKSRWQGNVFCGSMHEALHNYHTOKLSLSLSPCK 232  
DB 288 SDGSYFWYSLRVQKSTWGRGSLFACSVHEVLHNLTTTKTISRSLGK 335

Search completed: March 1, 2001, 09:17:41  
Job time: 259 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:13:22 ; Search time 79.26 seconds  
(without alignments)  
198.750 Million cell updates/sec

Title: US-09-389-782A-1  
Perfect score: 1263  
Sequence: 1 EPKSCDKTHTCPCPAPELL.....MHEALHNYTKLSLSPGK 232

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1263	100.0	330	1 GHHU	Ig gamma-1 chain C
2	1257	98.5	374	2 S69339	Ig heavy chain V r
3	1257	99.5	374	2 S72664	Ig gamma-1 chain C
4	1255	99.4	255	4 S31866	Ig gamma-1 chain C
5	1210	95.8	234	2 PT0207	Ig gamma-3 chain C
6	1176	93.1	377	2 A23511	Ig gamma-3 chain C
7	1174	93.0	377	2 A60764	Ig gamma-3 chain C
8	1151	91.1	289	1 G3HUW1	Ig gamma-3 heavy c
9	1145	90.7	326	1 G2HU	Ig gamma-2 chain C
10	1135	89.9	327	1 G4HU	Ig gamma-4 chain C
11	921	72.9	323	1 GHRB	Ig gamma chain C r
12	906.5	71.8	328	2 I47160	Ig gamma 2b chain
13	906.5	71.8	328	2 I47159	Ig gamma 2a chain
14	903	71.5	277	2 I47162	Ig gamma 4 chain c
15	896	70.9	329	1 G2GP	Ig gamma-2 chain c
16	885.5	70.1	328	2 I47158	Ig gamma 1 chain c
17	878.5	69.6	328	2 I47161	Ig gamma 3 chain c
18	856	67.8	470	2 S22080	Ig heavy chain pre
19	847.5	67.1	308	2 C30554	Ig heavy chain C r
20	847.5	67.1	472	2 S31459	Ig gamma-1 chain -
21	845.5	66.9	329	1 G3MSC	Ig gamma-3 chain C
22	842	66.7	333	2 PS0018	Ig gamma-2b chain
23	834.5	66.1	398	1 G3MSM	Ig gamma-3 chain C
24	832.5	65.9	444	2 PC4436	monoclonal antibod
25	822.5	65.1	324	1 G1MS	Ig gamma-1 chain C
26	822.5	65.1	326	2 PS0017	Ig gamma-1 chain C
27	817.5	64.7	393	1 G1MSM	Ig gamma-1 chain C
28	809.5	64.1	329	2 S00847	Ig gamma-2c chain
29	809	64.1	330	1 G2MSA	Ig gamma-2a chain

## ALIGNMENTS

RESULT 1  
GHHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433; MUID:82274238

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>

A:Cross-references: EMBL:Z17370

A:Note: this sequence has the Gln(17) allotypic marker, 97-Lys, and the Gln(1) marker

R:Harris, L.J.

submitted to the EMBL Data Library, October 1992

A:Reference number: S33904

A:Accession: S36861

A:Molecule type: DNA

A:Residues: 2-330 <HAR>

A:Cross-references: EMBL:Z17370

R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of

A:Reference number: S33887; MUID:83001943

A:Accession: S33887

A:Molecule type: DNA

A:Residues: 88-113;235-330 <TAK>

A:Cross-references: EMBL:Z17370

R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman

Biochemistry 9, 3161-3170, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq

A:Reference number: A90563; MUID:71064024

A:Contents: myeloma protein Eu

A:Accession: B90563

A:Molecule type: protein

A:Residues: 1-96, 'R', '98-135 <CUN>

A:Note: this sequence has the Gln(3) marker, 97-Arg

R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid seq

A:Reference number: A90564; MUID:71064025

A:Contents: Eu

A:Accession: A90564

A:Molecule type: protein

A:Residues: 136-154, 'Q', '156-165, 'Q', '178-194, 'N', '196-197, 'D', '199-238, 'E', '2

R:Postings, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976

A:Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni



C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

C:Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000

C:Accession: S31866

R:Filipula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene products.

A:Reference number: S31866

C:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FILL>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 99.4%; Score 1255; DB 4; Length 255;

Best Local Similarity 99.6%; Pred. No. 2.5e-90;

Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKF 60

Db 24 EKSCDKTHTCCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKF 83

QY 61 NWYVDGVEVHNAKTKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKT 120

Db 84 NWYVDGVEVHNAKTKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKT 143

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 180

Db 144 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 203

QY 181 PVLDSDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 232

Db 204 PVLDSDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 255

RESULT 5

PT0207

Ig gamma chain C region - chimpanzee

C:Species: Pan troglodytes (chimpanzee)

C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999

C:Accession: PT0207

R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.

Mol. Immunol. 28, 319-322, 1991.

A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.

A:Reference number: PT0207; MUID:91287716

A:Accession: PT0207

A:Molecule type: mRNA

A:Residues: 1-234 <EHR>

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:48-117/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 95.8%; Score 1210; DB 2; Length 234;

Matches 222; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKF 60

Db 10 EPKSCDTHTCCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKF 69

QY 61 NWYVDGVEVHNAKTKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKT 120

Db 70 NWYVDGVEVHNAKTKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKT 129

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 180

Db 130 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 189

QY 181 PVLDSDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 225

Db 190 PVLDSDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 234

RESULT 6

A23511

Ig gamma-3 chain C region (allotype G3m(b)) - human

C:Species: Homo sapiens (man)

C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999

C:Accession: A23511

R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.

Nucleic Acids Res. 14, 1779-1789, 1986

A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene:

A:Reference number: A23511; MUID:86148507

A:Accession: A23511

A:Molecule type: DNA

A:Residues: 1-377 <HUC>

A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056

C:Genetics:

A:Gene: GDB:IGHG3

A:Cross-references: GDB:119339; OMIM:147120

A:Map position: 14q32.33-14q32.33

A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 93.1%; Score 1176; DB 2; Length 377;

Best Local Similarity 92.7%; Pred. No. 5.6e-84;

Matches 215; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKF 60

Db 146 EPKSCDTHTCCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKF 205

QY 61 NWYVDGVEVHNAKTKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKT 120

Db 206 NWYVDGVEVHNAKTKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSNKALPAPIEKT 265

QY 121 ISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 180

Db 266 ISKTKGQPREPQVYTLPPSRDEMTKNOVSLTCLVKGFPSDIAVEVSNQGPENNYKTTTP 325

QY 181 PVLDSDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 232

Db 326 PVLDSDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPGK 377

RESULT 7

A60764

Ig gamma-3 chain C region, form LAT - human

C:Species: Homo sapiens (man)

C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 16-Jul-1999

C:Accession: A60764

R:Huck, S.; Lefranc, G.; Lefranc, M.P.

Immunogenetics 30, 250-257, 1989

A:Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 conv

A:Reference number: A60764; MUID:90007613

A:Accession: A60764

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-377 <HUC>

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:20-85/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 93.0%; Score 1174; DB 2; Length 377;

Matches 215; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 EPKSCDKTHTCCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKF 60

Db 146 EPKSCDTPPPCRPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 205  
Qy 61 NWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120  
Db 206 KWIYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 265  
Qy 121 ISKAGQPREPQVYTLPPSRDELAKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTP 180  
Db 266 ISKTKGQPREPQVYTLPPSRDELAKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTP 325  
Qy 181 PVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 232  
Db 326 PVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 377

RESULT 8  
G3HUI  
Ig gamma-3 heavy chain disease proteins - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1979 #sequence\_revision 23-Oct-1981 #text\_change 16-Jul-1999  
C:Accession: A90442; A92219; A90198; A93915; A02149  
R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.  
Biochemistry 19, 4304-4308, 1980  
A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-  
A:Reference number: A90442; MUID:81021548  
A:Contents: heavy chain disease protein Wis  
A:Accession: A90442  
A:Molecule type: protein  
A:Residues: 1-289 <FRA>  
A:Note: The molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain  
A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co  
A:Note: the sequence of residues 42-76 was taken from the reference that follows  
R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.  
J. Biol. Chem. 252, 883-889, 1977  
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication  
A:Reference number: A92219; MUID:77118561  
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein w  
A:Accession: A92219  
A:Molecule type: protein  
A:Residues: 12-97 <MTC>  
A:Note: The hinge region in gamma-3 chains is about four times as long as in other gamma  
idue segment (12-28)  
A:Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter  
R:Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.  
Biochem. Biophys. Res. Commun. 71, 907-914, 1976  
A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the  
A:Reference number: A90198; MUID:77021516  
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues  
A:Accession: A90198  
A:Molecule type: protein  
A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>  
A:Note: This protein lacks most of the V region, all of the CH1 region, and part of the  
R:Alexander, A.; Steinmetz, M.; Barricault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;  
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982  
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion  
A:Reference number: A93915; MUID:82247835  
A:Contents: heavy chain disease protein Omn  
A:Accession: A93915  
A:Molecule type: mRNA  
A:Residues: 12-70, 72-114, 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157  
A:Note: a carboxyl-terminal Lys is removed posttranslationally  
A:Note: this sequence may represent an allelic form or another gamma chain subclass  
C:Comment: The heavy chain disease protein Wis is shown.  
C:Genetics:  
A:Gene: GDB:IGHG3  
A:Cross-references: GDB:119339; OMIM:147120  
A:Map position: 14q32.33-14q32.33  
C:Superfamily: immunoglobulin C region; immunoglobulin; pyroglutamic acid  
C:Keywords: duplication; glycoprotein; immunoglobulin; immunoglobulin  
F:203-270/Domain: immunoglobulin homology <IM>  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 91.1% Score 1151; DB 1; Length 289;  
Best Local Similarity 90.5% Pred No. 3.5e-82;  
Matches 209; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

Qy 1 EPKSCDTHHTCCPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
Db 59 EPKSCDTPPPCRPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVQF 118  
Qy 61 NWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 120  
Db 119 KWIYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 178  
Qy 121 ISKAGQPREPQVYTLPPSRDELAKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTP 180  
Db 179 ISKTKGQPREPQVYTLPPSRDELAKNQVSLTCLVKGFYPSDIAVWESNGOPENNYKTP 238  
Qy 181 PVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 231  
Db 239 PVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 289

## RESULT 9

G2HU  
Ig gamma-2 chain C region - human  
C:Species: Homo sapiens (man)  
C:Date: 30-Apr-1981 #sequence\_revision 13-Jun-1983 #text\_change 21-Jul-2000  
C:Accession: A93906; A92809; A90752; A93132; A02148  
R:Ellison, J.; Hood, L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
A:Title: Linkage and sequence homology of two human immunoglobulin gamma chain  
A:Reference number: A93906; MUID:82197621  
A:Accession: A93906  
A:Molecule type: DNA  
A:Residues: 1-326 <ELL>  
A:Cross-references: GB:V00554; GB:J00230; NID:G32759; PIDN:CAB58438.1; PID:G6066056  
A:Note: Lys-326 is probably removed posttranslationally  
R:Wang, A.C.; Tung, E.; Fudenberg, H.H.  
J. Immunol. 125, 1048-1054, 1980  
A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, an  
A:Reference number: A92809; MUID:81007873  
A:Contents: myeloma protein Til  
A:Accession: A92809  
A:Molecule type: protein  
A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <WAN>  
A:Note: Trp-156 is at or near the complement-binding site  
R:Connell, G.E.; Parr, D.M.; Hofmann, T.  
Can. J. Biochem. 57, 758-767, 1979  
A:Title: The amino acid sequences of the three heavy chain constant region domains f  
A:Reference number: A90752; MUID:80001357  
A:Contents: myeloma protein Zie  
A:Accession: A90752  
A:Molecule type: protein  
A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, '222', 175, 'B', 177-193, 'D', 195-196, 'Q', 1  
A:Note: this sequence has since been revised  
R:Hofmann, T.; Parr, D.M.  
Mol. Immunol. 16, 923-925, 1979  
A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin  
A:Reference number: A93132; MUID:80114419  
A:Contents: Zie  
A:Accession: A93132  
A:Molecule type: protein  
A:Residues: 238-275 <HOF>  
R:Hofmann, T.; Parr, D.M.  
submitted to the Atlas, March 1980  
A:Reference number: A94591  
A:Contents: annotation: Zie, revisions to residues 25, 59, 60, and 264-268  
A:Note: the revised sequence differs from that shown in having 60 Ala and in the amid  
ned  
R:Milstein, C.; Frangione, B.  
Biochem. J. 121, 217-225, 1971  
A:Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.

A:Reference number: A90253; MUID:72033500  
 A:Contents: annotation; myeloma protein Sa, disulfide bonds  
 R:Frangione, B.; Milstein, C.; Pink, J.R.L.  
 Nature 221, 145-148, 1969  
 A:Title: Structural studies of immunoglobulin G.  
 A:Reference number: A93157; MUID:69064124  
 A:Contents: annotation; Sa, disulfide bonds  
 C:Genetics:  
 A:Gene: GDB:IGHG2  
 A:Cross-references: GDB:119338; OMIM:147110  
 A:Map position: 14q32.33-14q32.33  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger complexes.  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IM1>  
 F:133-202/Domain: immunoglobulin homology <IM2>  
 F:239-306/Domain: immunoglobulin homology <IM3>  
 F:14/Disulfide bonds: interchain (to light chain) #status experimental  
 F:27-83,140-200,246-304/Disulfide bonds: #status experimental  
 F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
 F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.7%; Score 1145; DB 1; Length 326;  
 Best Local Similarity 91.4%; Pred. No. 1.2e-81;  
 Matches 212; Conservative 9; Mismatches 7; Indels 4; Gaps 2;

QY 1 EPKSCDKTHCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 60  
 DB 99 ERKCCVE---CPCPAPP-VAGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 154  
 QY 61 NWYVDGVEVHNKTPREQYNSTYRVSVLTFLVHODWLNGLNGKEYCKVSNKALPAPIEKT 120  
 DB 155 NWYVDGVEVHNKTPREQYNSTYRVSVLTFLVHODWLNGLNGKEYCKVSNKALPAPIEKT 214  
 QY 121 ISKAKQPREPQYVTLPPSRDELTKNOVSLTCLVKGFPSDIAVEWESNGQPENNYKTTTP 180  
 DB 215 ISKTKGQPREPQYVTLPPSRDELTKNOVSLTCLVKGFPSDIAVEWESNGQPENNYKTTTP 274  
 QY 181 PVLDSGSEFLYSLKLVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 232  
 DB 275 PMLDSGSEFLYSLKLVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 326

RESULT 10  
 G4HU  
 Ig gamma-4 chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 16-Jul-1999  
 C:Accession: A90933; A90249; A02150  
 R:Ellison, J.; Buxbaum, J.; Hood, L.  
 DNA 1, 11-18, 1981  
 A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.  
 A:Reference number: A90933; MUID:83157104  
 A:Accession: A90933  
 A:Molecule type: DNA  
 A:Residues: 1-327 <ELL>  
 A:Note: the sequence was determined from the germline gene  
 R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.  
 Biochem. J. 117, 33-47, 1970  
 A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant region.  
 A:Reference number: A90249; MUID:70207560  
 A:Accession: A90249  
 A:Molecule type: protein  
 A:Residues: 1-30;81-326 <PIN>  
 C:Genetics:  
 A:Gene: GDB:IGHG4  
 A:Cross-references: GDB:119340; OMIM:147130  
 A:Map position: 14q32.33-14q32.33  
 A:Introns: 99/1; 111/1; 221/1  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger complexes.

C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IM1>  
 F:99-110/Region: hinge  
 F:134-203/Domain: immunoglobulin homology <IM2>  
 F:240-307/Domain: immunoglobulin homology <IM3>  
 F:14/Disulfide bonds: interchain (to light chain) #status experimental  
 F:27-83,141-201,247-305/Disulfide bonds: #status predicted  
 F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
 F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 89.9%; Score 1135; DB 1; Length 327;  
 Best Local Similarity 93.7%; Pred. No. 7.1e-81;  
 Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 11 CPCCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 70  
 DB 106 CPSCPAPEFLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVEVH 165  
 QY 71 NAKTKPREQYNSTYRVSVLTFLVHODWLNGLNGKEYCKVSNKALPAPIEKTISKAKGQPRE 130  
 DB 166 NAKTKPREQYNSTYRVSVLTFLVHODWLNGLNGKEYCKVSNKALPAPIEKTISKAKGQPRE 225  
 QY 131 PQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVEWESNGQPENNYKTTTPVLDSDGSFF 190  
 DB 226 PQVYTLPPSQEEMTKNOVSLTCLVKGFPSDIAVEWESNGQPENNYKTTTPVLDSDGSFF 285  
 QY 191 LYSKLVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 232  
 DB 286 LYSRLTVDSKRWQEGNVFCSCVMHEALHNHYTQKSLSLSPGK 327

RESULT 11  
 CHRB  
 Ig gamma chain C region - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 24-Apr-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Jul-1999  
 C:Accession: A91749; A90290; A93928; A90245; A94416; A02161  
 R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.  
 Immunogenetics 18, 387-397, 1983  
 A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I hap  
 A:Reference number: A91749; MUID:84030930  
 A:Accession: A91749  
 A:Molecule type: mRNA  
 A:Residues: 1-323 <BER>  
 A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-  
 R:Pratt, D.M.; Mole, L.E.  
 Biochem. J. 151, 337-349, 1975  
 A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunog  
 A:Reference number: A90290; MUID:76135469  
 A:Accession: A90290  
 A:Molecule type: protein  
 A:Residues: 1-47,'E',49-71,'PV',72-128 <PRA>  
 R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
 A:Title: Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma heavy ch  
 A:Reference number: A93928; MUID:83299917  
 A:Accession: A93928  
 A:Molecule type: mRNA  
 A:Residues: 88-103,'M',105-143,'E',145-184,'A',186,'E',188-266 <MAR>  
 A:Cross-references: GB:M16426; NID:g165111; PIDN:AAA31289.1; PID:g165112  
 A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic ma  
 R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.  
 Biochem. J. 116, 249-259, 1970  
 A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobul  
 A:Reference number: A90245; MUID:70110015  
 A:Accession: A90245  
 A:Molecule type: protein  
 A:Residues: 132-143,'E',145-161 <FRU>  
 R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
 In Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wikse  
 A:Reference number: A94416

QY 131 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNQ--PENNYKTTTPVLDSDGS 188

Best Local Similarity 72.18; Pred. No. 5.7e-63;

A;Contents: annotation; disulfide bonds  
A;Note: Cys-16 is involved in a heavy-light chain bond  
A;Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds  
C;Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.  
C;Complex: An immunoglobulin heterotrimer subunit consists of two identical light (lambda) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger multimers.  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin  
F;21-81/Domain: immunoglobulin homology <IM1>  
F;135-204/Domain: immunoglobulin homology <IM2>





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:41 ; Search time 40.97 Seconds  
(without alignments)  
316.083 Million cell updates/sec

Title: US-09-389-782A-2  
Perfect score: 2179  
Sequence: 1 MNKWLCALLVLDIIEWTT.....QKLFLEMIGNQVQSVKISCL 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	314	14.4	461	1	TNR2_HUMAN
2	312	14.3	474	1	TNR2_MOUSE
3	270	12.4	289	1	CD40_MOUSE
4	261	12.0	277	1	CD40_HUMAN
5	244	11.2	435	1	TNRC_HUMAN
6	233.5	10.7	325	1	VT2_SFVKA
7	228	10.5	269	1	CD40_BOVIN
8	220	10.1	326	1	VT2_MXVL
9	217	10.0	349	1	VC22_VARV
10	206	9.5	415	1	TNRC_MOUSE
11	199	9.1	427	1	NGFR_HUMAN
12	197.5	9.1	425	1	NGFR_RAT
13	190.5	8.7	323	1	FASA_BOVIN
14	184	8.4	416	1	NGFR_CHICK
15	182	8.4	335	1	FASA_HUMAN
16	177	8.1	256	1	41BB_MOUSE
17	176	8.1	332	1	FASA_PIG
18	167	7.7	255	1	41BB_HUMAN
19	165.5	7.6	461	1	TNR1_PIG
20	163	7.5	271	1	OX40_RAT
21	161.5	7.4	454	1	TNR1_MOUSE
22	160	7.3	455	1	TNR1_HUMAN
23	160	7.3	471	1	TNR1_BOVIN
24	153	7.0	461	1	TNR1_RAT
25	152.5	7.0	272	1	OX40_MOUSE
26	151	6.9	277	1	OX40_HUMAN
27	147.5	6.8	324	1	FASA_RAT
28	143	6.6	327	1	FASA_MOUSE
29	139	6.4	595	1	CD30_HUMAN
30	133	6.1	1786	1	LMB1_HUMAN
31	128.5	5.9	1790	1	LMB1_MOUSE
32	126.5	5.8	260	1	CD27_HUMAN
33	125	5.7	1680	1	FUR2_HUMAN
					P30432 drosophila

34 120.5 5.5 250 1 CD27\_MOUSE  
35 120.5 5.5 2911 1 FBN2\_HUMAN  
36 119.5 5.5 1639 1 LMGI\_DROME  
37 119 5.5 1609 1 LMGI\_HUMAN  
38 118 5.4 3712 1 LMA\_DROME  
39 116.5 5.3 915 1 PAC6\_MOUSE  
40 116 5.3 1955 1 AGRI\_CHICK  
41 116 5.3 3084 1 LMA1\_MOUSE  
42 116 5.3 3106 1 LMA2\_MOUSE  
43 115.5 5.3 713 1 TSA4\_GIALA  
44 115.5 5.3 1607 1 LMGI\_MOUSE  
45 115 5.3 417 1 WSL1\_HUMAN

P41272 mus musculus  
P35556 homo sapien  
P15215 drosophila  
P11047 homo sapien  
Q00174 drosophila  
Q04592 mus musculus  
P31696 gallus gall  
P19137 mus musculus  
Q06075 mus musculus  
P21849 giardia lam  
P02468 mus musculus  
Q93038 h wsl-1 pro

## ALIGNMENTS

RESULT 1  
ID TNR2\_HUMAN STANDARD; PRT; 461 AA.  
AC P20333;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR  
DE BINDING PROTEIN 2) (TBPII) (P80) (TNF-R2) (P75) (CD120B) (ETANERCEPT).  
GN TNFRSF1B OR TNFR2 OR TNFR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90260639; PubMed=2160731;  
RA Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,  
RA Dower S.K., Cosman D., Goodwin R.G.;  
RT "A receptor for tumor necrosis factor defines an unusual family of  
RT cellular and viral proteins.";  
RL Science 248:1019-1023(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91045991; PubMed=2172983;  
RA Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,  
RA Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;  
RT "A second tumor necrosis factor receptor gene product can shed a  
RT naturally occurring tumor necrosis factor inhibitor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96299745; PubMed=8661109;  
RA Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,  
RA Lepaslier D., Stallard B.J., Goeddel D.V., Desauvage F.J.,  
RA Brodeur G.M.;  
RT "Physical mapping and genomic structure of the human TNFR2 gene.";  
RL Genomics 35:94-100(1996).  
RN [4]  
RP SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=90349572; PubMed=2166946;  
RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,  
RA Ringold G.M.;  
RT "Complementary DNA cloning of a receptor for tumor necrosis factor  
RT and demonstration of a shed form of the receptor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).  
RN [5]  
RP SEQUENCE OF 27-31.  
RX MEDLINE=90110215; PubMed=2153136;  
RA Engelmann H., Novick D., Wallach D.;  
RT "Two tumor necrosis factor-binding proteins purified from human  
RT urine. Evidence for immunological cross-reactivity with cell surface  
RT tumor necrosis factor receptors.";  
RL J. Biol. Chem. 265:1531-1536(1990).  
RN [6]  
RP SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.  
RX MEDLINE=91056048; PubMed=2173696;





```

FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT CARBOHYD 153 153 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 289 AA; 32111 MW; C791CB6D2FEA574E CRC64;

Query Match 12.4%; Score 270; DB 1; Length 289;
Best Local Similarity 34.5%; Pred. No. 1.2e-12;
Matches 57; Conservative 26; Mismatches 58; Indels 14; Gaps 5;

QY 27 KYLHYDPETHQLKDKCAPGYLKQHCVRKTLVCPDHSYTDWHTSDEC---VYC 83
DB 30 QYLH-DGQ-----CDLCPGSRSLTSHCTALEKTCPCDSCGEFSAQNNREIRCHOHRHC 83
QY 84 SPVCKELQSVKQECNRTHNRVCECEGRYL---EIEFCLKHRSCTPPGSGVQVQAGTPERT 140
DB 84 EP--NQLRVKRGKGAESDVTCTCKEGHCTSKDCEACQAHTPCIPGFGVMEMATETTTDT 141
QY 141 VCKKCPDGFSSGTSKAPCIKHTNCSTFGLLLIQKGNATHDNVC 185
DB 142 VCHPCPVGFFSNQSLFEKCPYPTWSCDKNLEVLQKGTSTQNVIC 186

RESULT 4
CD40_HUMAN STANDARD; PRT; 277 AA.
AC P25942;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40)
DE (TUMOR NECROSIS FACTOR RECEPTOR 5).
GN TNFRSF5 OR CD40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89356608; PubMed=2475341;
RA Stamenkovic I., Clark E.A., Seed B.;
RT "A B-lymphocyte activation molecule related to the nerve growth
RT factor receptor and induced by cytokines in carcinomas."
RL EMBO J. 8:1403-1410(1989).
RN [2]
RP 3D-STRUCTURE MODELING OF 24-144.
RX MEDLINE=97189482; PubMed=9037712;
RA Bajorath J., Aruffo A.;
RT "Construction and analysis of a detailed three-dimensional model of
RT the ligand binding domain of the human B cell receptor CD40L."
RL Proteins 27:59-70(1997).
RN [3]
RP 3D-STRUCTURE MODELING OF 26-186 IN COMPLEX WITH CD40L.
RX MEDLINE=98266353; PubMed=9605317;
RA Singh J., Garber E., van Vlijmen H., Karpsusas M., Hsu Y.-M.,
RA Zheng Z., Naismith J.H., Thomas D.;
RT "The role of polar interactions in the molecular recognition of CD40L
RT with its receptor CD40."
RL Protein Sci. 7:1124-1135(1998).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD40 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd40.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60592; CAA43045.1;
DR PIR; S04460; S04460
DR PDB; ICDF; 01-APR-97.
DR MIM; 109535;
DR INTERPRO; IPR001368;
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
DR PROSITE; PS00500; TNFR_NGFR_2; 4.
KW Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal;
KW 3D-structure.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 277 CD40L RECEPTOR.
FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 POTENTIAL.
FT DOMAIN 216 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.
FT REPEAT 61 103 TNFR-CYS 1.
FT REPEAT 104 144 TNFR-CYS 2.
FT REPEAT 145 187 TNFR-CYS 3.
FT DISULFID 26 37 TNFR-CYS 4.
FT DISULFID 38 51
FT DISULFID 41 59
FT DISULFID 62 77
FT DISULFID 83 103
FT DISULFID 105 119
FT DISULFID 111 116
FT DISULFID 125 143
FT CARBOHYD 153 153
FT CARBOHYD 180 180
SQ SEQUENCE 277 AA; 30619 MW; BC8776EC2C4A5680 CRC64;

Query Match 12.0%; Score 261; DB 1; Length 277;
Best Local Similarity 33.9%; Pred. No. 5.1e-12;
Matches 58; Conservative 23; Mismatches 74; Indels 16; Gaps

QY 29 LHYDPETHG-----HQLLDKCAPGYLKQHCVRKTLVCPDHSYTDWHTSDEC 80
DB 18 VHEPPTACREKQYLLNSQCSLCPGKQKLVSDCTETETCLPGESEFLDTWNRTHC 77
QY 81 ---VYCSVPVCKELQSVKQECNRTHNRVCECEGRYL---EIEFCLKHRSCTPPGSGVQVQAG 134
DB 78 HQHKYCDPNLG--LRVQKGTSETDTICTCEGWHCTSEACSVLHRSCTPPGSGVQVQAG 135
QY 135 TPERNTVCKKCPDGFSSGTSKAPCIKHTNCSTFGLLLIQKGNATHDNVC 185
DB 136 TGVSDTICEPCPVGFFSNVSSAFERKCHPWTSCETKDLVVQVQAGTKNTDVC 186

RESULT 5
TNRC_HUMAN STANDARD; PRT; 435 AA.
AC P36941;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LYMPHOTXIN-BETA RECEPTOR (TUMOR NECROSIS FACTOR RECEPTOR
DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
GN LTR OR TNFR OR TNFRSF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=93252381; PubMed=8486360;
RA Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;
RT "Construction and evaluation of a hncDNA library of human l2p
RT transcribed sequences derived from a somatic cell hybrid."
RL Genomics 16:214-218(1993).

```

RN FUNCTION.  
 RP MEDLINE=94225209; PubMed=8171323;  
 RA Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,  
 RA Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;  
 RA "A lymphotoxin-beta-specific receptor.";  
 RL Science 264:707-710(1994).  
 CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN  
 CC IMMUNE DEVELOPMENT.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; L04270; AAA36757.1; -;  
 DR HSSP; P25942; 1CDF.  
 DR MIM; 600979;  
 DR INTERPRO; IPR001368; -;  
 DR PFAM; PF00020; TNFR\_C6; 4.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.  
 DR PROSITE; PS00500; TNFR\_NGFR\_2; 3.  
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.  
 FT SIGNAL 1 30 POTENTIAL.  
 FT CHAIN 31 435 LYMPHOTOXIN-BETA RECEPTOR.  
 FT DOMAIN 31 227 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 228 248 POTENTIAL.  
 FT DOMAIN 249 435 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 42 211 4 X TNFR-CYS.  
 FT REPEAT 42 81 TNFR-CYS 1.  
 FT REPEAT 82 124 TNFR-CYS 2.  
 FT REPEAT 125 168 TNFR-CYS 3.  
 FT REPEAT 169 211 TNFR-CYS 4.  
 FT DISULFID 43 58 BY SIMILARITY.  
 FT DISULFID 59 72 BY SIMILARITY.  
 FT DISULFID 62 80 BY SIMILARITY.  
 FT DISULFID 83 98 BY SIMILARITY.  
 FT DISULFID 101 116 BY SIMILARITY.  
 FT DISULFID 104 124 BY SIMILARITY.  
 FT DISULFID 126 132 BY SIMILARITY.  
 FT DISULFID 139 148 BY SIMILARITY.  
 FT DISULFID 142 167 BY SIMILARITY.  
 FT DISULFID 170 185 BY SIMILARITY.  
 FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 435 AA; 46709 MW; 624626B5022F656F CRC64;  
 Query Match 11.2%; Score 244; DB 1; Length 435;  
 Best Local Similarity 26.9%; Pred. No. 1.4e-10;  
 Matches 54; Conservative 32; Mismatches 93; Indels 22; Gaps 6;  
 QY 9 LVLVLLDIETWTQETLPP-----KYLHYDETGHQLCDKCAPGTYLKQHTVR 57  
 DB 18 VLGLFGLLAASQPAQVPPYASENQTCRDQKEYEYEQ--HRICCSRCPPGYYSKCSRI 75  
 QY 58 RKTLCVPCPDHSDWHTSDVCYSPCKELQSQKQ--ECNRTHNRVCEEGRY--- 112  
 DB 76 RDVVCATCAENSNEHNNYITICQLCRP-CDPVNGLEEETAPCTSKRKTQCRQGMFCFA 134  
 QY 113 --LEIEFLKHSRCPGSG--GVVQAGTPERTNVCKPDGFSGETSKAPCIKHTNCSTF 169  
 DB 135 WALECTHCELLSDCPGTEAELEKDEVGKGNHCVCKAGHFQNTSSPSARCQPHTRCENQ 194  
 QY 170 GLLLIQGNATHDNCVSGNRE 190  
 DB 195 GLVEAAPGTAQSDTTCKNPLE 215

RESULT 6  
 VT2\_SFVKA  
 ID VT2\_SFVKA STANDARD; PRT; 325 AA.  
 AC P25943;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).  
 GN T2.  
 OS Shope fibroma virus (strain Kasza) (SFV).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Leporipoxvirus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87321103; PubMed=2820128;  
 RA Upton C., Delange A.M., McFadden G.;  
 RT "Tumorigenic poxviruses: genomic organization and DNA sequence of the  
 RT telomeric region of the Shope fibroma virus genome.";  
 RL Virology 160:20-30(1987).  
 RN [2]  
 RP FUNCTION  
 RX MEDLINE=91207415; PubMed=1850261;  
 RA Smith C.A., Davis T., Wignall J.M., Din W.S., Farrah T., Upton C.,  
 RA McFadden G., Goodwin R.G.;  
 RT "T2 open reading frame from the Shope fibroma virus encodes a soluble  
 RT form of the TNF receptor.";  
 RL Biochem. Biophys. Res. Commun. 176:335-342(1991).  
 CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO  
 CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL  
 CC ANTIVIRAL EFFECTS OF THE CYTOKINE.  
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M17433; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; A23727; CAA01687.1; -;  
 DR PIR; B43692; B43692.  
 DR HSSP; P19438; 1TNR.  
 DR INTERPRO; IPR001368; -;  
 DR PFAM; PF00020; TNFR\_C6; 2.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.  
 DR PROSITE; PS00500; TNFR\_NGFR\_2; 1.  
 KW Receptor; Glycoprotein; Repeat; Signal.  
 FT SIGNAL 1 16 POTENTIAL.  
 FT CHAIN 17 325 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.  
 FT DOMAIN 27 186 4 X TNFR-CYS.  
 FT REPEAT 27 62 TNFR-CYS 1.  
 FT REPEAT 63 104 TNFR-CYS 2.  
 FT REPEAT 105 147 TNFR-CYS 3.  
 FT REPEAT 148 186 TNFR-CYS 4.  
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 325 AA; 35132 MW; 810530339198A71E CRC64;  
 Query Match 10.7%; Score 233.5; DB 1; Length 325;  
 Best Local Similarity 31.7%; Pred. No. 5.8e-10;  
 Matches 59; Conservative 19; Mismatches 91; Indels 17; Gaps 5;  
 QY 36 GHQ-----LLCDKCAPGTYLKQHTVRRKTLVPCPDHSDWHTSDVCYSPCKEL 90  
 DB 30 GHYDEKGLCCASCHPGFYASRLCGPGSNVTCSPEDGTFTASTNHAPACVSCRGCTGH 89  
 QY 91 QSVKQECNRTHNRVCEEGRYLEIE-----FCLKHSRCPGSGVVQAGTPERTNVCKK 144

```
Db 90 LSEQPCDRTDRVNCSTGNYCLLKQNGCRICAPQTKCPAGY-VSGHTRAGDTLCEK 148
QY 145 CPDGFSGSETSKAPCIKHTKNSFGLLLIOKNATHDNCVSGNREATQKCGIDVTL--- 201
Db 149 CPPIHTYDLSPTERCSTGTSFNYISVGFNLYPV-NETSCTTTAGHNEVIKTKFTVTLNYT 207
QY 202 -CEEAF 206
Db 208 DCDPVF 213

RESULT 7
CD40_BOVIN STANDARD; PRT; 269 AA.
AC Q28203;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (FRAGMENT).
GN TNFRSF5 OR CD40.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97281252; PubMed=9135560;
RA Hirano A., Brown W.C., Estes D.M.;
RT "Cloning, expression and biological function of the bovine CD40
RT homologue: role in B-lymphocyte growth and differentiation in
RT cattle.";
RL Immunology 90:294-300(1997).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U57745; AAC48710.1;
CC HSP; P25942; ICDF
CC INTERPRO: IPR001368;
CC PFAM; PF00020; TNFR_C6; 4.
CC PROSITE; PS00652; TNFR_NGFR_1; 1.
CC PROSITE; PS50050; TNFR_NGFR_2; 1.
CC Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal.
CC SIGNAL 1 19
CC CHAIN 20 >269 CD40L RECEPTOR.
CC DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 194 215
CC DOMAIN 216 >269 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 215 187 4 X TNFR-CYS.
CC REPEAT 25 60 TNFR-CYS 1.
CC REPEAT 61 103 TNFR-CYS 2.
CC REPEAT 104 144 TNFR-CYS 3.
CC REPEAT 145 187 TNFR-CYS 4.
CC CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC NON_TER 269 269
CC SEQUENCE 269 AA; 29983 MW; 746903F30F95F387 CRC64;

Query Match 10.5%; Score 228; DB 1; Length 269;
Best Local Similarity 31.9%; Pred. No. 1.le-09;
Matches 60; Conservative

QY 29 LHYDPETG-----HOLLCDKCAPGTLYLKQHCTVRRKTLVCPDPHSYDTSWHTSDEC 80
```

```
Db 18 VHSEPATACGEKQYPVNSLCCDLPPGQKLVNDCTEVSKTEQSCGKGFEFLSTNRKVC 77
QY 81 ---VYCSP-VCKELQSVQKBCNRRHNRVCEBGRYL---EIFELCKHRSPPGSGVQVA 133
Db 78 HEHRYCNFNGLRIQS---EGTLNTDTTCVCEGQHCSTHTCESCTPHSLCLUPFGVAKI 134
QY 134 GTPERTVCKKCPDGFSGSETSSKAPCIKHTKNSFGLLLIOKNATHDNCVSGNREATQ 193
Db 135 ATGLLDTVCECPGLGFFSNVSSAFEKCHRWTSCKRGLVLEQHVGNKNTDVCVCGESRMT 194
QY 194 KCGIDVTL 201
Db 195 LWIPIVTM 202

RESULT 8
VT2_MYXVL STANDARD; PRT; 326 AA.
AC P29825;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN 72).
GN TN2.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=91335768; PubMed=1651597;
RA Upton C., Macen J.L., Schreiber M., McPadden G.;
RT "Myxoma virus expresses a secreted protein with homology to the tumor
RT necrosis factor receptor gene family that contributes to viral
RT virulence.";
RL Virology 184:370-382(1991).
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M95181; AAA46632.1;
CC EMBL; A23729; CAA01688.1;
CC PIR; A40566; GQVZML.
CC HSP; P19438; ITNR.
CC INTERPRO: IPR001368;
CC PFAM; PF00020; TNFR_C6; 2.
CC PROSITE; PS00652; TNFR_NGFR_1; 2.
CC PROSITE; PS50050; TNFR_NGFR_2; 2.
CC Receptor; Glycoprotein; Repeat; Signal.
CC SIGNAL 1 16
CC CHAIN 17 326 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.
CC DOMAIN 27 186 4 X TNFR-CYS.
CC REPEAT 27 62 TNFR-CYS 1.
CC REPEAT 63 104 TNFR-CYS 2.
CC REPEAT 105 147 TNFR-CYS 3.
CC REPEAT 148 186 TNFR-CYS 4.
CC CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 326 AA; 35208 MW; ABBF027E947292FF CRC64;

Query Match 10.1%; Score 220; DB 1; Length 326;
```

```

Best Local Similarity 29.1%; Pred. No. 5.4e-09;
Matches 53; Conservative 24; Mismatches 85; Indels 20; Gaps 5;

QY 39 LDCDKCAPGYLKHCHTCTVRRKTLCPVCPDHSYTDWHTSDCVVCPVKELQSVKQECN 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 LCCTSCPPGSYARLCPGSDTVCSPCKNTFTASTNHAPACVSCRGCRGTGHLSESQSCD 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 99 RTHNRVCECEGRYLEIE-----FCLKHRSPPGSGVQAGTPERTVCKKCPDGFSSG 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 KTRDRVCDCSAGNYCLLKQEGCRICAPKTKCPAGYG-VSGHTRTGDVLCTKCPRTYSD 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 153 ETSSKAPCIKHTNCSFGLLLIQG-----NATHDNCVSGNREATQKGDIVTL-----CEE 204
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 AVSSTETC-----TSSFNYISVFENLYPVNDTCTTTAGPNEVVKTSFVTLNHTDCDP 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 205 AF 206
Db 212 VF 213

RESULT 9
VC22_VARV STANDARD; PRT; 349 AA.
ID VC22_VARV
AC P34015;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
GN G4R.
OS Variola virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INDIA-1967 / ISOLATE IND3;
RX MEDLINE=9320281; PubMed=8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
RL protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
CC -1- SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69198; CAA49137.1; -
CC DR EMBL; X67117; CAA47540.1; -
CC DR PIR; D36858; D36858.
CC DR PIR; S35987; S35987.
CC DR PIR; S46888; S46888.
CC DR HSP; P19438; INCF.
CC INTERPRO: IPR001368; -.
CC PFAM: PF00020; TNFR_c6; 2.
CC PROSITE: PS00652; TNFR_NGFR_1; 2.
CC PROSITE: PS00050; TNFR_NGFR_2; 2.
CC Repeat.
KW DOMAIN 31 108 2 X TNFR-CYS.
FT REPEAT 31 66 TNFR-CYS 1.
FT REPEAT 67 108 TNFR-CYS 2.
SQ SEQUENCE 349 AA; 38189 MW; D45D40B5C6E780EF CRC64;

Query Match
Best Local Similarity 10.0%; Score 217; DB 1; Length 349;
Matches 92; Conservative 51; Mismatches 169; Indels 84; Gaps 17;

QY 4 WLCCALVLDIIETWTOETLPKYLHYDPE-TGHQLLCDKCAPGYLKHCHTCTVRRKTL 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 11 FLSCIIINGRAAPYT-----PPNGKCKTEYKRNHLCCLSCPPGYASRLCDSKTNTQC 65
QY 63 VPDPDSYTDWHTSDCVVCPVKELQSVKQECNTHNRVCECEGRYLEI-----E 116
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 TPGSGTFTSRNNHLPACLSGRCNQNQVETSCNTHNRICECSPGYYCLLKSGSGCK 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 117 FCLKHRSPPGSGVQAGTPERTVCKKCPDGFSSGTSKAPCIKHTNCSFGLLLIQ- 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 ACVSQTKCGIGYG-VSGHTSVDGVICSPCGFYSHTVSSADKCEPVPN-NTFNVIDVEI 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 176 KGNATHDNCVSGNREATQKGDIVTLCEEAFFRAVPTKLIIPNLWSVLVDSLPGTKVNAE 235
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 TLYPVNDTCT-----RTTTGLSE-----SILTSEL----- 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 236 SVERIKRRHSS-----QEOTFQLLK-----LMKHQRDQEMVKKIITQIDDLCESSVQRH 284
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 ----TIYNNHTDCNPVFREEYFSLNKVATSGFFTGEMYQNISK-----VCTLNFEIK 260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 285 LGHSNLTTEQLLALMESLPCKISPEIERTRKTKCSSEQLLKLLSLWRINKGDDTLKG 344
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 CNKGSFSKQLTRAKND--DGMMSHSETVTLAGDCLSSVDIYILYSNTNAQDYETDIS- 317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 345 LMYALKHL--KTSHP-----KTVTHSLRKTMRFL 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 --YRVGNVLDDSHMPCGSCNIHKPITNS--KPTRFL 349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
TNRC_MOUSE STANDARD; PRT; 415 AA.
ID TNRC_MOUSE
AC P50284;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
GN LTBR OR TNFCR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CVB; TISSUE=LUNG;
RX MEDLINE=96072804; PubMed=7594541;
RA Force W.R., Walter B.N., Hession C., Tizard R., Kozak C.A.,
RA Browning J.L., Ware C.F.;
RT "Mouse lymphotoxin-beta receptor. Molecular genetics, ligand binding,
RT and expression.";
RL J. Immunol. 155:5280-5288(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96163885; PubMed=8586432;
RA Nakamura T., Tashiro K., Nazarea M., Nakano T., Sasayama S.,
RA Honjo T.;
RT "The murine lymphotoxin-beta receptor cDNA: Isolation by the signal
RT sequence trap and chromosomal mapping.";
RL Genomics 30:312-319(1995).
CC -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC IMMUNE DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U29173; AAA68964.1; -.
CC DR EMBL; L38423; AAB00846.1; -.
CC DR EMBL; U30798; AAA81334.1; -.
CC DR HSP; P25942; 1CDF.

```









Qy	149	FFSGEYSSKAPCIKHNCSTFGLLLI-----QKGNATHD--NYVCSGNREAT	199
Dd	164	--GSRS-----HAN-SLWALLILLPIVLIIYKVKSERNKKNDYCNSAASNDEGR	212
Qy	193	QKCGIDVTICEAAFFFAFPTKIIPNWLSVLVDSLPGTK----VNAESVERIK-----	241
Dd	213	QLNLTDVDL-----GKYIPS-----IAEQMRITEVKFVRKNGMEEAKIDDIMHD	257
Qy	242	RRISSOEQTFOILLKLW--KHONRD-----OEMVKKIIODI-DLCSSVORHLGHSNLT	292
Dd	258	NVHETAEQKVOLLRNYYQSHGKNAYCTITKSLPKALAEKICDIVMKDITNERENANLN	317
Qy	293	E 293	
Dd	318	E 318	
 RESULT 14			
ID	NGFR_CHICK		
ID	NGFR_CHKCK	STANDARD; PRT; 416 AA.	
AC	P18519;		
DT	01-NOV-1990 (Rel. 16, Created)		
DT	01-NOV-1990 (Rel. 16, Last sequence update)		
DT	01-OCT-2000 (Rel. 40, Last annotation update)		
DE	LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR) (GP80-LNGFR) (P75 ICD).		
GN	NGFR		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
RC	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BRAIN;		
RX	MEDLINE=90166579; PubMed=2560385;		
RA	Large T.H., Weskamp G., Helder J.C., Radeke M.J., Misko T.P.,		
RA	Shooter E.M., Reichardt L.F.;		
RT	"Structure and developmental expression of the nerve growth factor		
RT	receptor in the chicken central nervous system.";		
RL	Neuron 2:1123-1134(1989).		
RN	[2]		
RC	SEQUENCE OF 21-416 FROM N.A.		
RX	MEDLINE=90152140; PubMed=2154393;		
RA	Heuer J.G., Fatemie-Nainie S., Wheeler E.F., Bothwell M.;		
RT	"Structure and developmental expression of the chicken NGF receptor."		
RT	Dev. Biol. 137:287-304(1990).		
CC	-!- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,		
CC	NT-3, AND NT-4.		
CC	-!- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE		
CC	BOND FORMATION.		
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.		
CC	-!- PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.		
CC	-!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.		
CC	-!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.		
DR	PIR; JN0006; JN0006.		
DR	PIR; A60504; A60504.		
DR	HSP; P07174; INCR.		
DR	INTERPRO; IPR000488; .		
DR	INTERPRO; IPR001368; .		
DR	PFAM; PF00020; TNFR_c6; 4.		
DR	PFAM; PF00531; death; 1.		
DR	PROSITE; PS00652; TNFR_NGFR_1; 3.		
DR	PROSITE; PS00050; TNFR_NGFR_2; 3.		
DR	PROSITE; PS50017; DEATH_DOMAIN; 1.		
KW	Receptor; Neurogenesis; Transmembrane; Glycoprotein; Repeat;		
KW	Phosphorylation; Signal.		
FT	SIGNAL	1 19	
FT	CHAIN	20 416	
FT			LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR.
FT	DOMAIN	29 239	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	240 261	POTENTIAL.
FT	DOMAIN	262 416	CYTOPASMIC (POTENTIAL).
FT	DOMAIN	23 181	4 X TNFR-CYS.

```

FT REPEAT 23 57 TNFR-CYS 1.
FT REPEAT 58 100 TNFR-CYS 2.
FT REPEAT 101 140 TNFR-CYS 3.
FT REPEAT 141 181 TNFR-CYS 4.
FT DOMAIN 188 236 SER/THR-RICH.
FT DOMAIN 333 410 DEATH DOMAIN.
FT DISULFID 24 35 BY SIMILARITY.
FT DISULFID 36 49 BY SIMILARITY.
FT DISULFID 59 56 BY SIMILARITY.
FT DISULFID 59 75 BY SIMILARITY.
FT DISULFID 78 91 BY SIMILARITY.
FT DISULFID 81 99 BY SIMILARITY.
FT DISULFID 101 114 BY SIMILARITY.
FT DISULFID 117 130 BY SIMILARITY.
FT DISULFID 120 138 BY SIMILARITY.
FT DISULFID 141 156 BY SIMILARITY.
FT DISULFID 159 172 BY SIMILARITY.
FT DISULFID 162 180 BY SIMILARITY.
FT CARBOHYD 52 52 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 36 36 C -> Y (IN REF. 2).
FT CONFLICT 173 173 T -> K (IN REF. 2).
FT CONFLICT 276 276 N -> S (IN REF. 2).
FT CONFLICT 396 396 K -> R (IN REF. 2).
SQ SEQUENCE 416 AA; 44654 MW; 6BCEAAB54F4D2D56 CRC64;

Query Match 8.4%; Score 184; DB 1; Length 416;
Best Local Similarity 30.9%; Pred. No. 2.7e-06;
Matches 46; Conservative 28; Mismatches 67; Indels 8; Gaps 6;

QY 41 CDKCAPFTYKQKHTVRRKTLVCPDPH-SYTDSWHTSDCYVCSVPYKELQSVKQECNR 99
DB 36 CKACNLGEGVVPQGV-NQIVCEPCLDSVTYSVATEPCPKCTQ-CVGLHSMSPACVE 93
QY 100 THNRVCECEGRYLETEF---CLKHRSCPPGSGVVGQAGTPERTVCKKCPDGFSGTSS 156
DB 94 SDDAVCRCAYG-YFQDELSSCKECSCEVGFGLMFPFCDSDQTVCECEPGTFSDEANF 152
QY 157 KAPCIKHTNGSTFGLLLIOGNATHDNVC 185
DB 153 VDCPLCTICEE-NEVMVKCTATSDAEC 180

RESULT 15
FASA_HUMAN STANDARD; PRT; 335 AA.
AC P25445;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE (APO-1 ANTIGEN) (CD95 ANTIGEN).
GN TNFRSF6 OR APT1 OR FAS OR FASL.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE-91309137; PubMed-1713127;
RA Itoh N., Yonehara S., Ishii A., Yonehara M., Mizushima S.I.,
RA Sameshima M., Hase A., Seto Y., Nagata S.;
RT "The polypeptide encoded by the cDNA for human cell surface antigen
RL Fas can mediate apoptosis.";
RL Cell 66:233-243(1991).
[2]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 226-240; 269-291 AND 321-335.
RX MEDLINE-92268122; PubMed-1375228;
RA Oehm A., Behrmann I., Falk W., Pawlita M., Maier G., Klas C.,
RA Li-Weber M., Richards S., Dhein J., Trauth B.C., Ponstingl H.,
RA Kramer P.H.;
RT "Purification and molecular cloning of the APO-1 cell surface
RT antigen, a member of the tumor necrosis factor/nerve growth factor
RT receptor superfamily. Sequence identity with the Fas antigen.";
```

```

J. Biol. Chem. 267:10709-10715(1992).
[3]
RN STRUCTURE BY NMR OF 218-335.
RX MEDLINE-97122332; PubMed-8967952;
RA Huang B., Eberstadt M., Olejniczak E.T., Meadows R.P., Fesik S.W.;
RT "NMR structure and mutagenesis of the Fas (APO-1/CD95) death domain.";
RL Nature 384:638-641(1996).
CC -|- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED
CC SUICIDE OF MATURE T-CELLS, OR BOTH.
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -|- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC -|- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -|- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -|- DATABASE: NAME=PROW; NOTE=CD guide CD95 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd95.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M57454; AAA63174.1; -
DR EMBL; X63717; CAA45250.1; -
DR PIR; A40036; A40036.
DR PIR; S24543; S24543.
DR PDB; 1DDF; 12-NOV-97.
DR MIM; 134637; -
DR INTERPRO; IPR000488; -
DR INTERPRO; IPR001388; -
DR PFAM; PF00020; TNFR_c6; 2.
DR PFAM; PF00531; death; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal;
KW 3d-structure.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 335 FASL RECEPTOR.
FT DOMAIN 17 173 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 174 190 POTENTIAL.
FT DOMAIN 191 335 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 47 166 3 X TNFR-CYS.
FT REPEAT 47 83 TNFR-CYS 1.
FT REPEAT 84 127 TNFR-CYS 2.
FT REPEAT 128 166 TNFR-CYS 3.
FT DOMAIN 230 314 DEATH DOMAIN.
FT CARBOHYD 118 118 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 335 AA; 37732 MW; 0139942535111410 CRC64;

Query Match 8.4%; Score 182; DB 1; Length 335;
Best Local Similarity 25.5%; Pred. No. 2.9e-06;
Matches 76; Conservative 34; Mismatches 114; Indels 74; Gaps 15;

QY 19 TQETLPPKYLHYDPETGHQLLCKCAPGYLKHQHTVR-RKTLVCPDP-HSYTDSWHT 76
DB 42 TVTETQNLGLHHDGQFCHK----PCPPGERKARDCTVNGDEPDCVPCQSGKETDKAHF 97
QY 77 SDECVYCSVPYKELQSVKQE--CNRTNRYVCECEEGRYLEIFELCKHRSCPPGSGVVGQAG 134
DB 98 SSKCRRCR-LCDEGHGLEVEINCTRTQNTKCRCKPNFCNSTVC---EHCDP----- 145
```

Search completed: March 1, 2001, 09:17:44  
Job time: 262 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:42 ; Search time 79.26 Seconds  
(without alignments)  
343.530 Million cell updates/sec

Title: US-09-389-782A-2  
Perfect score: 2179  
Sequence: 1 MNKWLCALLVLDIIEWTT.....OKLFLEMIGNOVQVKISCL 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues  
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	314	14.4	459	2 I48854	gene murine tumour
2	312	14.3	461	1 A35356	tumor necrosis fac
3	312	14.3	474	2 B38634	tumor necrosis fac
4	270	12.4	289	2 A46515	B cell-associated
5	270	12.4	305	2 A46476	B cell-associated
6	261	12.0	277	2 A60771	B-cell activation
7	244	11.2	435	2 I54182	tumor necrosis fac
8	233	10.7	325	2 B43692	T2 protein - rabbi
9	220	10.1	326	1 GQVZML	T2 protein - myxom
10	217	10.0	349	2 D36858	gene G4R protein -
11	216	9.9	349	2 D72175	G2R protein - vari
12	215	9.9	348	2 T28623	hypothetical prote
13	199	9.1	427	1 GQHUN	nerve growth facto
14	197	9.1	425	1 A26431	nerve growth facto
15	184	8.4	416	1 JN0006	nerve growth facto
16	182	8.4	335	2 A40036	apoptosis-mediati
17	179	8.2	314	2 I37383	FAS soluble protei
18	177	8.1	256	2 B32393	T-cell antigen 4-1
19	167	7.7	255	2 I38426	4-1BB - human
20	166	7.6	255	2 JT0752	lymphocyte activat
21	165	7.6	461	2 JC4302	tumor necrosis fac
22	163	7.5	271	2 SI2783	Ox40 antigen precu
23	161	7.4	454	1 GQMT1	tumor necrosis fac
24	161	7.4	454	2 I57826	tumor necrosis fac
25	160	7.3	455	1 GQHUT1	tumor necrosis fac
26	153	7.0	461	1 GQRUT1	tumor necrosis fac
27	152	7.0	272	2 I48700	gene ox40 protein
28	151	6.9	277	2 I37552	Ox40 homolog - hum
29	147	6.8	324	2 JC2395	Fas antigen precu

30	143	6.6	327	2 A46484	apoptosis-mediati
31	142	6.5	2823	2 T23064	hypothetical prote
32	142	6.5	3102	2 T43291	laminin alpha chai
33	139	6.4	595	2 A42086	CD30 antigen precu
34	137	6.3	899	2 G02428	subtilisin-like pr
35	137	6.3	915	2 JC6148	subtilisin-like pr
36	133	6.1	1786	1 MMHUB1	laminin beta-1 cha
37	131	6.0	260	1 A46517	CD27 antigen precu
38	128	5.9	786	2 A48456	cocyst wall protei
39	128	5.9	1252	2 S36016	cocyst wall protei
40	128	5.9	1790	1 MMFEB1	laminin beta-1 cha
41	126	5.8	1548	2 S34583	serine proteinase
42	125	5.7	1680	2 A43434	furin (EC 3.4.21.7
43	122	5.6	1372	2 T25933	hypothetical prote
44	120	5.5	250	1 A49053	CD27 antigen precu
45	120	5.5	2918	2 A54105	fibrillin-2 precu

ALIGNMENTS

RESULT 1  
I48854  
gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
C:Accession: I48854  
R:Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.  
Mamm. Genome 5, 726-727, 1994  
A:Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.  
A:Reference number: I48854; MUID:95178848  
A:Accession: I48854  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-459 <RES>  
A:Cross-references: EMBL:X76401; NID:9433830; PIDN:CAA53981.1; PID:9433831  
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology  
F:151-188/Domain: NGF receptor repeat homology <NGP>

Query Match 14.4%; Score 314; DB 2; Length 459;  
Best Local Similarity 34.6%; Pred. No. 1.7e-14;  
Matches 66; Conservative 20; Mismatches 79; Indels 26; Gaps 4;

QY	18	WTQETLPPKYL--HYDPETGH-----QLLCKCAPGYLKHCHVRRKTLQ	62
DB	2	WAGHTVPAQVVLTPYKPEPEGCISQEIYDRKAQMCACPPGQYVKKHFCNATSDTVCI	61
QY	63	VPCPDHSYTDSWHTSDCVYCSPVKELQSVKQECNTRNVRVCECEGRYLEIEF----	117
DB	62	ADCEASMYTQWNQFRCLSCSSCSTDQVETRACTKQONRVCAACEAGRYCALKTHSGSC	121
QY	118	--CLKHRCPPGSGVQAGTPERNVTKCPDGFSSGETSSKAPCIKHNTSGLLLIQ	175
DB	122	RQMRLSKCGPGFGVASSRAPNGNVLCACAPGTFSDTSDTCDVRPHRICS----ILAI	177
QY	176	KGNAHDNVCS	186
DB	178	PGNASTDVAVCA	188

RESULT 2  
A35356  
tumor necrosis factor receptor type 2 precursor - human  
N:Alternate names: 75k tumor necrosis factor receptor  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094  
R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dover, Science 248, 1019-1023, 1990  
A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular  
A:Reference number: A35356; MUID:90260639  
A:Accession: A35356

A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-461 <SMI>  
A:Cross-references: GB:M32315; NID:g189185; PIDN:AAA59929.1; PID:g189186  
R:Kolomo, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squires, Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990  
A:Title: A second tumor necrosis factor receptor gene product can shed a naturally occurring  
A:Reference number: A36475; MUID:91045991  
A:Accession: A36475  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-195, 'R', 197-461 <KOH>  
A:Cross-references: GB:M55994; GB:M38549; NID:g339757; PIDN:AAA36755.1; PID:g339758  
R:Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, M. Cytokine 2, 231-237, 1990  
A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular, A:Reference number: A48416; MUID:91370690  
A:Accession: A48416  
A:Status: preliminary  
A:Molecule type: mRNA; protein  
A:Residues: 23-461 <DEM>  
A:Cross-references: GB:S63358; NID:g235648; PIDN:AB119824.1; PID:g235649  
A:Note: sequence extracted from NCBI backbone (NCBIN:63368, NCBI:P:63371)  
R:Heller, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M. Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990  
A:Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demonstration of its expression in human cells  
A:Reference number: A36007; MUID:90349572  
A:Accession: A36007  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 116-140, 'P', 142-195, 'R', 197-362, 'T', 364-461 <HEL>  
A:Cross-references: GB:M35857; NID:g339751; PIDN:AAA63262.1; PID:g339752  
R:Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus, M. J. Biol. Chem. 265, 20131-20138, 1990  
A:Title: Purification and partial amino acid sequence analysis of two distinct tumor necrosis factor receptors  
A:Reference number: A23666; MUID:91056048  
A:Accession: A23666  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 23-40, 65-69, 136-141, 300-306 <LOE>  
R:Engelmann, H.; Novick, D.; Wallach, D. J. Biol. Chem. 265, 1531-1536, 1990  
A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Evidence of their interaction with the TNF receptor  
A:Reference number: A35010; MUID:90110215  
A:Accession: B35010  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 27-31 <ENG>  
R:Kuhnert, P.; Kemper, O.; Wallach, D. Gene 150, 381-386, 1994  
A:Title: Cloning, sequencing and partial functional characterization of the 5' region of the human TNF receptor  
A:Reference number: I38094; MUID:95121934  
A:Accession: I38094  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-37 <RES>  
A:Cross-references: EMBL:X80021; NID:g666044; PIDN:CAA56324.1; PID:g825701  
C:Genetics:  
A:Gene: GDB:TNFR2  
A:Cross-references: GDB:125914; OMIM:191191  
A:Map position: lp36.2-1p36.2  
A:Introns: 26/3  
A:Note: the list of introns is incomplete  
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology  
C:Keywords: duplication; glycoprotein; receptor; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-416/Product: tumor necrosis factor receptor type 2 #status experimental <MAT>  
F:40-76/Domain: NGF receptor repeat homology <NG1>  
F:78-119/Domain: NGF receptor repeat homology <NG2>  
F:120-162/Domain: NGF receptor repeat homology <NG3>  
F:164-201/Domain: NGF receptor repeat homology <NG4>  
F:262-279/Domain: transmembrane #status predicted <TMN>  
F:280-461/Domain: intracellular #status predicted <INT>

F:171,193/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 14.38; Score 312; DB 1; Length 461;  
Best Local Similarity 34.78; Pred. No. 2.4e-14;  
Matches 69; Conservative 19; Mismatches 85; Indels 26; Gaps 5;  
Qy 8 ALLVLLDIIEWTQTETLPPK--YLHYDPEIGH-----QLLCDKCAPGYLKQH 53  
Db 9 ALAVGLEL--WAAHALPAQVAFPTPAPEPGSTCRLREYDQTAQMCCSKSPGGHAKVF 56  
Qy 54 CTVRRKTLVPCPDHSHYSDSHWTSKQVYCVCKELQSVKQECNRTNHRVCEEGRYL 113  
Db 67 CTKTSDTVCDSCEDSTYTLQNNWVPECLSGSCSSDQVETQACTREQNRICTCRPGWVC 126  
Qy 114 EI-----EFLCKHRSCPPGSGVQVAGTPERTVCKKCPDGGFFSGTSSKAPCIKHTNCS 167  
Db 127 ALSKQEGCRLCAPLRKCRPGFVARPGTETSDVVKCAPGTFSTNTSTDICRPHQICN 186  
Qy 168 TFGLLLIQKGNATHDNVCS 186  
Db 187 VVAI-----PGNASMDAVCT 201  
RESULT 3  
B38634  
tumor necrosis factor receptor type 2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 23-Jul-1999  
C:Accession: B38634; A40254; S54816  
R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen, Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991  
A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis factor receptors  
A:Reference number: A38634; MUID:91187885  
A:Accession: B38634  
A:Molecule type: mRNA  
A:Residues: 1-474 <LEW>  
A:Cross-references: GB:M60459; NID:g199827; PIDN:AAA39752.1; PID:g199828  
R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; J. Mol. Cell. Biol. 11, 3020-3026, 1991  
A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors for tumor necrosis factor  
A:Reference number: A40254; MUID:91246168  
A:Accession: A40254  
A:Molecule type: mRNA  
A:Residues: 1-474 <GOO>  
A:Cross-references: GB:M60459; NID:g199827; PIDN:AAA39752.1; PID:g199828  
R:Kisssonerghis, M.; Fallowes, R.; Feldmann, M.; Chernajovsky, Y. submitted to the EMBL Data Library, May 1995  
A:Description: Characterization of the promoter region of the murine p75-TNF receptor  
A:Reference number: S54816  
A:Accession: S54816  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-22 <KIS>  
A:Cross-references: EMBL:X87128; NID:g809043; PIDN:CAA0618.1; PID:g809044  
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology  
C:Keywords: cytokine receptor; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-474/Product: tumor necrosis factor receptor type 2 #status predicted <MAT>  
F:40-77/Domain: NGF receptor repeat homology <NG1>  
F:79-120/Domain: NGF receptor repeat homology <NG2>  
F:166-203/Domain: NGF receptor repeat homology <NG4>

Query Match 14.38; Score 312; DB 2; Length 474;  
Best Local Similarity 34.38; Pred. No. 2.4e-14;  
Matches 70; Conservative 22; Mismatches 84; Indels 28; Gaps 5;  
Qy 5 LCCALLVLLDIIEWTQTETLPPKYL--HYDPEIGH-----QLLCDKCAPGY 49  
Db 6 LWVALVFEQL--WATGHTVPAQVVLTPYKPEPGVCEQISQEIYDRKAQMCCKAPPGY 63  
Qy 50 LKQHCTVRRKTLVPCPDHSHWTSKQVYCVCKELQSVKQECNRTNHRVCECEE 109

Db 64 VKHFNKTSVTCADCEASMYTQVNFRTCLSCSSCTTDOVEIRACTQQRNVCACEA 123  
QY 110 GYLLEIEF-----CLKHRSPPGSGVOAGTPERNVCKKPDGFFSGTSSKAPCIK 162  
Db 124 GRVCAKTHSGSCRCQRLSKGPGFVASSRAPNGVNLKACAPGFTSDTSSDVCRP 183  
QY 163 HTNCSTFGLLLIQGNATHDNYCS 186  
Db 184 HRICS-----ILAIPGNASTDAVCA 203

## RESULT 4

A64615  
B cell-associated surface molecule CD40, short splice form - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 26-May-2000  
C:Accession: A46515  
R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockayne, J. Immunol. 149, 3921-3926, 1992  
A:Title: Genomic structure and chromosomal mapping of the murine CD40 gene.  
A:Reference number: A46515; MUID:93094586  
A:Accession: A46515  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-289 <GRI>  
A:Cross-references: GB:M83312; NID:g1553058; PIDN:AAB08705.1; PID:g1553059; GB:M94126; N  
A:Experimental source: BALB/C, liver  
A:Note: sequence extracted from NCBI backbone (NCBIP:120357)  
C:Comment: For an alternative splice form, see PIR:A46476  
C:Superfamily: CD27 antigen; NGF receptor repeat homology  
C:Keywords: alternative splicing  
F:105-144/Domain: NGF receptor repeat homology <NGF>

Query Match 12.4%; Score 270; DB 2; Length 289;  
Best Local Similarity 34.5%; Pred. No. 1.2e-11;  
Matches 57; Conservative 26; Mismatches 68; Indels 14; Gaps 5;

QY 27 KYLHYDPETGHQLLCKCAPGTYLKHQCTVRRKTLVPCPDHSTYDSWHTSDEC---VYC 83  
Db 30 QYLH-DGQ-----CCDLQCPGSRULTSHTALEKTQCHPCDSGEFSAQWNRREIRCHQHRHC 83  
QY 84 SPVCKELQSVKQECNRTHNRVCECEGRYL---EIFECLKHRSPPGSGVVGAGTPERT 140  
Db 84 EP--NOGLRVKKEGTAEVDVCTCKEGQCHTSKDCEACAOHTPCIPGFGVEMATETDT 141  
QY 141 VCKKCPDGFSGTSSKAPCIKHTNCTFGLLLIQGNATHDNYC 185  
Db 142 VCHPCPVGFTSNQSSLFKCYPTWTSCEDKNLEVLQKGTSTQTNVIC 186

## RESULT 5

A46476  
B cell-associated surface molecule CD40, long splice form - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 18-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 26-May-2000  
C:Accession: A46476  
R:Torres, R.M.; Clark, E.A.  
J. Immunol. 148, 620-626, 1992  
A:Title: Differential increase of an alternatively polyadenylated mRNA species of murine  
A:Reference number: A46476; MUID:92105763  
A:Accession: A46476  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-305 <TOR>  
A:Cross-references: GB:M83312; NID:g1553058  
A:Note: sequence extracted from NCBI backbone (NCBIN:75206, NCBIP:75207)  
C:Comment: this translation is not annotated in GenBank entry M83312, release 113.0  
C:Superfamily: CD27 antigen; NGF receptor repeat homology  
C:Keywords: alternative splicing; transmembrane protein  
F:105-144/Domain: NGF receptor repeat homology <NGF>

Query Match 12.4%; Score 270; DB 2; Length 305;  
Best Local Similarity 34.5%; Pred. No. 1.2e-11;  
Matches 57; Conservative 26; Mismatches 68; Indels 14; Gaps 5;

QY 27 KYLHYDPETGHQLLCKCAPGTYLKHQCTVRRKTLVPCPDHSTYDSWHTSDEC---VYC 83  
Db 30 QYLH-DGQ-----CCDLQCPGSRULTSHTALEKTQCHPCDSGEFSAQWNRREIRCHQHRHC 83  
QY 84 SPVCKELQSVKQECNRTHNRVCECEGRYL---EIFECLKHRSPPGSGVVGAGTPERT 140  
Db 84 EP--NOGLRVKKEGTAEVDVCTCKEGQCHTSKDCEACAOHTPCIPGFGVEMATETDT 141  
QY 141 VCKKCPDGFSGTSSKAPCIKHTNCTFGLLLIQGNATHDNYC 185  
Db 142 VCHPCPVGFTSNQSSLFKCYPTWTSCEDKNLEVLQKGTSTQTNVIC 186

## RESULT 6

A60771  
B-cell activation protein CD40 precursor - human  
N:Alternate names: B-cell surface antigen Bp50  
C:Species: Homo sapiens (man)  
C:Date: 03-Jun-1993 #sequence\_revision 03-Feb-1994 #text\_change 21-Jul-2000  
C:Accession: S04460; A60771  
R:Stamenkovic, I.; Clark, E.A.; Seed, B.  
EMBO J. 8, 1403-1410, 1989  
A:Title: A B-lymphocyte activation molecule related to the nerve growth factor receptor  
A:Reference number: S04460; MUID:89356608  
A:Accession: S04460  
A:Molecule type: mRNA  
A:Residues: 1-277 <STA>  
A:Cross-references: EMBL:X60592; NID:g29850; PIDN:CAA43045.1; PID:g29851  
R:Brasch-Andersen, S.; Paulie, S.; Kono, H.; Nika, H.; Aspenstroem, P.; Perlmann, P.  
J. Immunol. 142, 562-567, 1989  
A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-  
A:Reference number: A60771; MUID:89093941  
A:Accession: A60771  
A:Molecule type: protein  
A:Residues: 21-50 <BRA>  
A:Experimental source: Burkitt lymphoma cell line Raji  
C:Genetics:  
A:Gene: GDB:CD40  
A:Cross-references: GDB:215268; OMIM:109535  
A:Map Position: 20q12-20q13.2  
C:Superfamily: CD27 antigen; NGF receptor repeat homology  
C:Keywords: B-cell; glycoprotein; phosphoprotein; surface antigen; transmembrane prot  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-277/Product: B-cell activation protein CD40 #status experimental <MAT>  
F:21-193/Domain: extracellular #status predicted <EXT>  
F:194-215/Domain: transmembrane #status predicted <TM>  
F:216-277/Domain: intracellular #status predicted <CYT>  
F:153,180/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 12.0%; Score 261; DB 2; Length 277;  
Best Local Similarity 33.9%; Pred. No. 4.6e-11;  
Matches 58; Conservative 23; Mismatches 74; Indels 16; Gaps 4;

QY 29 LHYDPETG-----HOLLCDKCAPGTYLKHQCTVRRKTLVPCPDHSTYDSWHTSDEC 80  
Db 18 VHEPPTACREKQYVLINSQCCLSCOPGKLVSDCTETFTETCLPGESEFLDTWNRTHC 77  
QY 81 ---VYCSPVCKELQSVKQECNRTHNRVCECEGRYL---EIFECLKHRSPPGSGVVGAG 134  
Db 78 HOHRYCDPNLGL--LRVQOKGTSETDICTCEBWHCTSEACESCVLHRSCTSPGFGVQKQIA 135  
QY 135 TPERNTVCKKCPDGFSGTSSKAPCIKHTNCTFGLLLIQGNATHDNYC 185  
Db 136 TGVSDTICEPCPVGFTSNQSSLFKCYPTWTSCEDKNLEVLQKGTSTQTNVIC 186

RESULT 7  
154182  
tumor necrosis factor receptor 2-related protein - human  
C:Species: Homo sapiens (man)  
C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 17-Mar-2000  
C:Accession: 154182  
R:Baens, M.; Chaffanet, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.  
Genomics 16, 214-218, 1993  
A:Title: Construction and evaluation of a hncDNA library of human 12p transcribed sequences  
A:Reference number: 154182; MUID:93252381  
A:Accession: 154182  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-435 <RES>  
A:Cross-references: GB:L04270; NID:g339761; PIDN:AAA36757.1; PID:g339762  
C:Genetics:  
A:Gene: GDB:LTBR  
A:Cross-references: GDB:1230195; OMIM:600979  
A:Map position: 12p13.3-12p13.1  
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

Query Match 11.2%; Score 244; DB 2; Length 435;  
Best Local Similarity 26.9%; Pred. No. 1.1e-09;  
Matches 54; Conservative 32; Mismatches 93; Indels 22; Gaps 6;

QY 9 LVLGLDIIETWTTQETLPP-----KYLHYDPETGHQLLCKAPGVTLKQKCTVR 57  
DB 18 VLGLFGLLAASQVAPVYASENQTCROEKEYEYEPQ--HRICCSRCPPGTVSAKCSRI 75  
QY 58 RKTLCVPCPDHSDYTSWHTSDVCYVCKELQSVKQ--ECNRTHNRVCEEGRY--- 112  
DB 76 RDTVCATCAENSYNEHNYLTICQLCRP--CDPVMGLEETAPCTSKRTKTCRCQCPGMFCAA 134  
QY 113 --LETEFLKLRSCPPGS--GVVQAGTPERTNVCKKCPDGFSGETSSKAPCKTHNCSTF 169  
DB 135 WALECTHCELLSDCPPTGAELKDEVGNNHCVPCKAGHFONTSSPACQPHTRCENQ 194  
QY 170 GLLLIQKGNATHDNVCSGNRE 190  
DB 195 GLVEAAPGTAQSDTTCKNPLE 215

RESULT 8  
B43692  
T2 protein - rabbit fibroma virus  
C:Species: rabbit fibroma virus, Shope fibroma virus  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: B43692  
R:Upton, C.; DeLange, A.M.; McFadden, G.  
Virology 160, 20-30, 1987  
A:Title: Tumorigenic poxviruses: genomic organization and DNA sequence of the telomeric  
A:Reference number: B43692; MUID:87321103  
A:Accession: B43692  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-325 <UPT>  
A:Cross-references: GB:M17433  
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology  
F:64-105/Domain: NGF receptor repeat homology <NG2>  
F:106-147/Domain: NGF receptor repeat homology <NG3>

Query Match 10.7%; Score 233.5; DB 2; Length 325;  
Best Local Similarity 31.7%; Pred. No. 4.4e-09;  
Matches 59; Conservative 19; Mismatches 91; Indels 17; Gaps 5;

QY 36 GHQ-----LLCKCAPGYLKHCTVRRKTLVCPDHSYTSWHTSDVCYVCKEL 90  
DB 30 GHDEKQDGLCCASCHPGFYASRLCGPGSNTVCSPCEDGTFTASTNHAPACVSCRCGPCTGH 89  
QY 91 QSVKQECNRTHNRVCEEGRYLEIE-----FCLKHSRCPGSGVVOAGTPERTNVCKK 144

DB 90 LSESPQCDRTHRVNCSTGNVCLLKGGCRICAPQTKCPAGYG-VSGHTRAGDTLCEK 148  
QY 145 CPDGFSGETSSKAPCIKHTNCSTFGLLLIQKGNATHNVCSGNREATQKCGIDVTL--- 201  
DB 149 CPPHTYSLSLSPTRCGTSFNVISVGFNLVYPV-NETSCSTTTAGHNEVIKTEFTVLTNYT 207  
QY 202 -CEEAF 206  
DB 208 DCDPVF 213

RESULT 9  
GQVZML  
T2 protein - myxoma virus (strain Lausanne)  
C:Species: myxoma virus  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 18-Jun-1999  
C:Accession: A40566  
R:Upton, C.; Macen, J.L.; Schreiber, M.; McFadden, G.  
Virology 184, 370-382, 1991  
A:Title: Myxoma virus expresses a secreted protein with homology to the tumor necrosis  
A:Reference number: A40566; MUID:91335768  
A:Accession: A40566  
A:Molecule type: DNA  
A:Residues: 1-326 <UPT>  
A:Cross-references: GB:M95181; GB:M37976; NID:g332309; PIDN:AAA46632.1; PID:g332310  
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology  
C:Keywords: glycoprotein  
F:64-105/Domain: NGF receptor repeat homology <NG2>  
F:106-147/Domain: NGF receptor repeat homology <NG3>  
F:66,181,205,238/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.1%; Score 220; DB 1; Length 326;  
Best Local Similarity 29.1%; Pred. No. 3.8e-08;  
Matches 53; Conservative 24; Mismatches 85; Indels 20; Gaps 5;

QY 39 LLCKCAPGYLKHCTVRRKTLVCPDHSYTSWHTSDVCYVCKELQVQKQCN 98  
DB 38 LCCTSCPPGYSVASRLCGPGSDTVCSFCKNETFTASTNHAPACVSCRCRGTGHLSESQSCD 97  
QY 99 RTHNRVCEEGRYLEIE-----FCLKHSRCPGSGVVOAGTPERTNVCKKCPDGFSG 152  
DB 98 KTRDRVCDCSAGNYCLLKGGCRICAPKTKCPAGYG-VSGHTRTGDVLTCKCPRYTSD 156  
QY 153 ETSSKAPCIKHTNCSTFGLLLIQK-----NATHDNVCSGNREATQKCGIDVTL---CEE 204  
DB 157 AVSSTETC-----TSSFNVISVEFNLYPVNDTSCSTTTAGPNEVVKTSFSTLHTDCDP 211  
QY 205 AF 206  
DB 212 VF 213

RESULT 10  
D36858  
gene G4R protein - variola virus  
N:Alternate names: B28R protein (COP)  
C:Species: variola virus  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 08-Oct-1999  
C:Accession: D36858; S46888; S32385; S35987  
R:Blinov, V.M.  
submitted to GenBank, November 1992  
A:Description: not shown.  
A:Reference number: A36859  
A:Accession: D36858  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-349 <BLI>  
A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49137.1; PID:g457087  
A:Experimental source: strain India-1967, ssp. major, isolate Ind3  
R:Kolykhalov, A.A.; Blinov, V.M.; Gytarov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; F  
submitted to the EMBL Data Library, April 1992  
A:Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H





F;109-147/Domain: NGF receptor repeat homology <NG3>  
F;149-189/Domain: NGF receptor repeat homology <NG4>  
F;197-248/Region: serine/threonine-rich  
F;251-272/Domain: transmembrane #status predicted <TRM>  
F;273-427/Domain: intracellular #status predicted <INT>  
F;60/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 199; DB 1: Length 427;  
Best Local Similarity 31.7%; Pred. No. 1.4e-06;  
Matches 52; Conservative 25; Mismatches 71; Indels 16; Gaps 6;

QY 9 LLVLLDIIEWTQTETLPPKYLHYDPETGHOLLCKCAPGYTLKHCHTKLVRRKTLVCPCPDH 68  
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 17 LLLLLGVSLGAKAECAPTGLYTHSGE-----CCKACNLGGEVAQPCGA-NQTVECPCLDS 70  
| | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 69 -SYTDSWHTSDECVYSPVKCELOSVKQECNRNHNRVCEEGRYLE-----TEFLCKHR 122  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 71 VTFSDDVSATPECKPCTE-CVGLQSAPCAVEADAVCRCAVGYYQDETTCRCEAC---R 126  
| | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 123 SCPFGSGVQAAGPERNTWCCKCPDGFSSGETSKAPCIKHTNC 166  
| | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 127 VCEAGSLVFSCQDKNTVCEECDPGTYSDGANHVDCPLCTVC 170  
| | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 14

A26431  
nerve growth factor receptor precursor, low affinity - rat  
N:Alternate names: NGF receptor  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A26431; PH1229  
R:Radeke, M.J.; Misko, T.P.; Hsu, C.; Herzenberg, L.A.; Shooter, E.M.  
Nature 325, 593-597, 1987  
A>Title: Gene transfer and molecular cloning of the rat nerve growth factor receptor.  
A:Reference number: A26431; MUID:87115859  
A:Accession: A26431  
A:Molecule type: mRNA  
A:Residues: 1-425 <RAD>  
A:Cross-references: GB:X05137; NID:56755; PIDN:CAA28783.1; PID:g56756  
R:Metzlis, M.; Timmusk, T.; Allikmeets, R.; Saarma, M.; Persson, H.  
Gene 121, 247-254, 1992  
A>Title: Regulatory elements and transcriptional regulation by testosterone and retin  
A:Reference number: PH1229; MUID:93077038  
A:Accession: PH1229  
A:Molecule type: DNA  
A:Residues: 1-20 <MET>  
A:Cross-references: GB:X61269  
C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastom  
C:Comment: The cysteine-rich region of the extracellular domain may form part or all  
C:Comment: This protein is thought to form a high-affinity receptor when it associate  
C:Genetics:  
A:Introns: 20/3  
C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology  
F;1-29/Domain: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor  
F;30-425/Product: nerve growth factor receptor #status predicted <MAT>  
F;30-425/Product: nerve growth factor receptor #status predicted <EXT>  
F;30-251/Domain: extracellular #status predicted <EXT>  
F;33-66/Domain: NGF receptor repeat homology <NG1>  
F;68-109/Domain: NGF receptor repeat homology <NG2>  
F;110-148/Domain: NGF receptor repeat homology <NG3>  
F;150-190/Domain: NGF receptor repeat homology <NG4>  
F;198-249/Region: serine/threonine-rich  
F;252-273/Domain: transmembrane #status predicted <MEM>  
F;274-425/Domain: intracellular #status predicted <INT>  
F;61/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 197.5; DB 1: Length 425;  
Best Local Similarity 22.1%; Pred. No. 1.8e-06;  
Matches 95; Conservative 59; Mismatches 152; Indels 123; Gaps 19;

OV 9 LLVLLDIIEWTQTETLPPKYLHYDPETGHOLLCKCAPGYTLKHCHTKLVRRKTLVCPCPDH 68

```

18 LLLILGVSSGGAKETCTGLYTHSGE-----CCKACNLGEGVAQPCGA-NQTVCEPCLDN 71
69 -SYTDSWHTSDECVYCSVPVCKELQSVKQCNTRTHNRVCECEBGRYLEIE--FCLKHRSCLP 125
72 VTFSDVVSATEPCKPCTE--CLGLQSMSAPCVRADDVACRCAYGYODEETGHCACSVCE 130
126 PGSGVVQAGTPERNVTCKKCPDGFSGETSSKAPCIKHTNCSTFGLLLIQKGNATHDNVC 185
131 VSGSLVFCQDKQNTVCEBCEPCTYSDEANHVDPCLPCTVCEDETERQL----- 178
186 SGNREATQKCGIDVTLCEBAFFR-----AVPTK--IIPNWSVL 223
179 ---RECTPWADAE---CEBIPGRWIPRSTPEGSDSTAPSTOEPPVPEQDLVPSTVADM 232
224 VDSLPGCTKYNABVERIKRHSQE-----QTFQLLKLWK--HQNRQOE 265
233 VTTVMG-----SSQPVVTRGTDNLIPVCSILAADVGLVAYIAFKRWNSCKQKQGA 286
266 MVKKIIQ-----DIDLCESSVQRHLGHSNLTTEQLA-----LMESLPCKK-- 306
287 NSRPVNTPTPEGEKLSHDSGISVDSQSLHDQOHTQTASGQALKGDNLYSSLPLTKRE 346
307 -----ISPETIER-TRKTKSSEQLKLKLSLWRKNGDQD--TLKG 344
347 EVEKLLNGDTWRHLAGELGYQPEHIDFTHEAC---PVRALLASW---GAQDSATLDA 398
345 LMVALKHLK 353
399 LLAALRRIQ 407
```

RESULT 15

```

JN0006
nerve growth factor receptor, low affinity precursor - chicken
N:Alternate names: NGF receptor
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JN0006; A60504
R:Large, T.H.; Weskamp, G.; Helder, J.C.; Radeke, M.J.; Misko, T.P.; Shooter, E.M.; Reichman, R.; Heuer, J.G.; Fatemie-Nainie, S.; Wheeler, E.F.; Bothwell, M.
Dev. Biol. 137, 287-304, 1990
A:Title: Structure and developmental expression of the nerve growth factor receptor in the chicken
A:Reference number: JN0006; MUID:90166579
A:Accession: JN0006
A:Molecule type: mRNA
A:Residues: 1-416 <LAR>
A:Experimental source: embryonic chick brain
R:Heuer, J.G.; Fatemie-Nainie, S.; Wheeler, E.F.; Bothwell, M.
Dev. Biol. 137, 287-304, 1990
A:Title: Structure and developmental expression of the chicken NGF receptor.
A:Reference number: A60504; MUID:90152140
A:Accession: A60504
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 21-35,'Y',37-172,'K',174-275,'S',277-395,'R',397-416 <HEU>
C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma cells, and on the cysteine-rich region of the extracellular domain may form part or all of the binding site.
C:Comment: This protein is thought to form a high-affinity receptor when it associates with the nerve growth factor receptor; NGF receptor repeat homology
C:Superfamily: nerve growth factor receptor; heterodimer; monomer; phosphoprotein; receptor; tyrosine kinase
C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor; tyrosine kinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-416/Product: nerve growth factor receptor #status predicted <MAT>
F:21-239/Domain: extracellular #status predicted <EXT>
F:24-57/Domain: NGF receptor repeat homology <NG1>
F:59-100/Domain: NGF receptor repeat homology <NG2>
F:101-139/Domain: NGF receptor repeat homology <NG3>
F:141-181/Domain: NGF receptor repeat homology <NG4>
F:189-237/Region: serine/threonine-rich
F:240-261/Domain: transmembrane #status predicted <MEM>
F:262-416/Domain: intracellular #status predicted <INT>
F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```

Query Match 8.4%; Score 184; DB 1; Length 416;
Best Local Similarity 30.9%; Pred. No. 1.5e-05;
Matches 46; Conservative 28; Mismatches 67; Indels 8; Gaps 6;
QY 41 CDKCAPGTLYKQHCIVRRKTLVPCPDH-SYTDSWHTSDECVYCSVPVCKELQSVKQECNR 99
Db 36 CRACNLGEGVQPCGV-NQTVCEPCLDSVTYSQTVSATEPCKPCTQ-CVGLHMSAPCVE 93
QY 100 THNRVCECEBGRYLEIEF---CLKHRSCTPGSGVVQAGTPERNVTCKKCPDGFSGETSS 156
Db 94 SDAVCRCAYG-YFQDELSGSKCEICEVGFGLMFPFPCRDQTVCECEPCTVCEDEANF 152
QY 157 KAPCIKHTNCSTFGLLLIQKGNATHDNVC 185
Db 153 VDPCLPCTICEE-NEVMVKECTATSDAEC 180
Search completed: March 1, 2001, 09:15:45
Job time: 143 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:20:05 ; Search time 135.68 seconds  
(without alignments)  
345.542 Million cell updates/sec

Title: US-09-389-782A-5  
Perfect score: 2240  
Sequence: 1 ETFPPKYLHYDEETSHQLLC.....VMHEALHNHYTKSLSPG 400

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_15:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1013.5	45.2	401	4 Q00300	O00300 homo sapien
2	1013	45.2	372	4 Q0UHP4	Q0UHP4 homo sapien
3	893.5	39.9	401	11 O08712	O08712 mus musculus
4	877	39.2	401	11 O08727	O08727 rattus norv
5	827	36.9	437	11 Q9RIA4	Q9RIA4 mus musculus
6	433	19.3	300	4 Q95407	Q95407 homo sapien
7	405.5	18.1	302	13 Q9PUS0	Q9pus0 salvelinus
8	345	15.4	459	11 Q62327	Q62327 mus musculus
9	343.5	15.3	439	4 Q16042	Q16042 homo sapien
10	338	15.1	482	11 O08734	O08734 mus musculus
11	297	13.3	655	4 Q75509	Q75509 homo sapien
12	280.5	12.5	684	13 Q90544	Q90544 ginglymosto
13	277.5	12.4	384	4 Q9UP60	Q9up60 homo sapien
14	273.5	12.2	416	4 Q9NEP6	Q9nep6 homo sapien
15	258.5	11.5	616	4 Q9Y606	Q9y606 homo sapien
16	256.5	11.5	625	11 Q35305	Q35305 mus musculus
17	241	10.8	349	12 O57099	O57099 monkeypox v
18	239	10.7	349	12 O57291	O57291 monkeypox v
19	239	10.7	349	12 O57100	O57100 monkeypox v

#### ALIGNMENTS

RESULT 1

O00300

ID	O00300	PRELIMINARY;	PRT;	401 AA.
AC	O00300;	O60236;		
DT	01-JUL-1997	(TrEMBLrel. 04, Created)		
DT	01-JUL-1997	(TrEMBLrel. 04, Last sequence update)		
DT	01-OCT-2000	(TrEMBLrel. 15, Last annotation update)		
DE	OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)			
DE	(OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11b).			
GN	TNFRSF11B OR OPG OR OCIF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
RN	TISSUE=KIDNEY;			
RC	MEDLINE=97262071; PubMed=9108485;			
RX	Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,			
RA	Luethy R., Nguyen H.O., Woodson S., Bennett L., Boone T., Shimamoto G.,			
RA	Derose M., Elliott R., Colombero A., Tan H.-L., Traill G., Sullivan J.,			
RA	Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,			
RA	Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,			
RA	Suggs S., Boyle W.J.;			
RT	*Osteoprotegerin: a novel secreted protein involved in the regulation			
RT	of bone density.";			
RL	Cell 89:309-319(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LUNG FIBROBLAST;			
RX	MEDLINE=98151033; PubMed=9492069;			
RA	Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,			
RA	Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,			
RA	Tsuda E., Morinaga T., Higashio K.;			
RT	*Identity of osteoclastogenesis inhibitory factor (OCIF) and			
RT	osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits			
RT	osteoclastogenesis in vitro."			
RL	Endocrinology 139:1329-1337(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RN	TISSUE=PLACENTA;			
RC	MEDLINE=98351569; PubMed=9688283;			

057101 monkeypox v  
057102 monkeypox v  
057277 monkeypox v  
057103 monkeypox v  
057108 monkeypox v  
057112 variola vir  
Q85407 variola vir  
057284 camelpox vi  
057098 camelpox vi  
057110 variola vir  
057111 variola vir  
Q89118 variola vir  
Q89098 variola vir  
057116 cowpox viru  
057097 camelpox vi  
057123 cowpox viru  
057308 cowpox viru  
057109 variola vir  
057305 cowpox viru  
057122 cowpox viru  
057118 cowpox viru  
057120 cowpox viru  
073559 cowpox viru  
057117 cowpox viru  
057115 cowpox viru  
057092 ectromelia

20 239 10.7 349 12 057101  
21 239 10.7 349 12 057102  
22 235.5 10.5 348 12 057277  
23 235.5 10.5 348 12 057103  
24 235.5 10.5 348 12 057108  
25 229 10.2 348 12 057112  
26 229 10.2 348 12 057112  
27 226.5 10.1 349 12 057284  
28 226.5 10.1 349 12 057098  
29 226 10.1 349 12 057110  
30 226 10.1 349 12 057111  
31 226 10.1 349 12 057118  
32 226 10.1 349 12 057118  
33 223 10.0 350 12 057116  
34 222.5 9.9 349 12 057097  
35 222.5 9.9 350 12 057123  
36 222.5 9.9 355 12 057109  
37 221 9.9 349 12 057109  
38 220.5 9.8 349 12 057305  
39 219 9.8 326 12 057122  
40 218.5 9.8 360 12 057118  
41 217.5 9.7 326 12 057120  
42 217.5 9.7 351 12 073559  
43 216.5 9.7 351 12 057117  
44 212.5 9.5 347 12 057115  
45 210.5 9.4 316 12 057092

RA Moringa T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;  
RT "Cloning and characterization of the gene encoding human  
RT osteoprotegerin/osteoclastogenesis-inhibitory factor";  
RL Eur. J. Biochem. 254:685-691(1998).  
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY  
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY  
CC SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,  
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN  
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN  
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.  
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
DR EMBL; AB002146; BAA25910.1; -.  
DR EMBL; AB008822; BAA32076.1; -.  
DR EMBL; AB008821; BAA32076.1; JOINED.  
DR EMBL; U94332; AAB53709.1; -.  
DR HSSP; P25942; 1CDF.  
DR MIM; 602643; -.  
DR INTERPRO; IPR001368; -.  
DR PFAM; PF00020; TNFR\_C6; 3.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
DR PROSITE; PS00500; TNFR\_NGFR\_2; 2.  
DR PRODOM; PD00071; -; 1.  
KW Glycoprotein; Repeat; Cytokine; Signal.  
FT SIGNAL 1 21  
FT CHAIN 1 21 BY SIMILARITY.  
FT DOMAIN 22 401 OSTEOPROTEGERIN.  
FT REPEAT 23 183 4 X TNFR-CYS.  
FT REPEAT 23 63 TNFR-CYS 1.  
FT REPEAT 64 106 TNFR-CYS 2.  
FT REPEAT 107 143 TNFR-CYS 3.  
FT REPEAT 144 201 TNFR-CYS 4.  
FT DOMAIN 306 365 DEATH DOMAIN.  
FT DISULFID 41 54 BY SIMILARITY.  
FT DISULFID 44 62 BY SIMILARITY.  
FT DISULFID 65 80 BY SIMILARITY.  
FT DISULFID 83 97 BY SIMILARITY.  
FT DISULFID 87 105 BY SIMILARITY.  
FT DISULFID 118 142 BY SIMILARITY.  
FT DISULFID 145 160 BY SIMILARITY.  
FT CARBOHYD 98 98 POTENTIAL.  
FT CARBOHYD 152 152 POTENTIAL.  
FT CARBOHYD 165 165 POTENTIAL.  
FT CARBOHYD 178 178 POTENTIAL.  
FT CARBOHYD 289 289 POTENTIAL.  
FT CONFLICT 263 263 A -> D (IN REF. 2 AND 3).  
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 45.2%; Score 1013.5; DB 4; Length 401;  
Best Local Similarity 68.7%; Pred. No. 3.1e-81;  
Matches 193; Conservative 12; Mismatches 37; Indels 39; Gaps 5;

Qy 1 ETFFPKYLHYDEETSHQLLCKDKCPGTYLKQCTAKWTKVACAPCPDHYTDSWHTSDECL 60  
Db 22 ETFFPKYLHYDEETSHQLLCKDKCPGTYLKQCTAKWTKVACAPCPDHYTDSWHTSDECL 81  
Qy 61 YCSPVCKELQYVQECNTRNHRVCECKEGRYLEIEFCLKLRHSCPPGFGVQAGTPERNTV 120  
Db 82 YCSPVCKELQYVQECNTRNHRVCECKEGRYLEIEFCLKLRHSCPPGFGVQAGTPERNTV 141  
Qy 121 CKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTOKVDKTHTC 180  
Db 142 CKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTOK 194  
Qy 181 PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 240  
Db 195 --CGIDVTLCCEAFAFRFAVPTK-----FTPNWLSVLVD-----NLPCTKVA 234

Qy 241 AKTKPREQYNSTYRVSVLTVLHODWLNKGYKCKVSNKA 281  
Db 235 ESVERIKRQHSSQEOTFOLLKL-----W-----KHQNK 263

## RESULT 2

Q9UHP4 PRELIMINARY; PRT; 372 AA.  
AC Q9UHP4;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE OSTEOPROTEGERIN (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;  
RL "Cloning and Expression of Osteoprotegerin from Homo sapiens";  
DR EMBL; AF134187; AAF20168.1; -.  
DR HSSP; P25942; 1CDF.  
DR INTERPRO; IPR001368; -.  
DR PFAM; PF00020; TNFR\_C6; 3.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
DR PROSITE; PS00500; TNFR\_NGFR\_2; 2.  
FT NON\_TER 1  
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 45.2%; Score 1013; DB 4; Length 372;  
Best Local Similarity 73.4%; Pred. No. 3.2e-81;  
Matches 190; Conservative 10; Mismatches 35; Indels 24; Gaps 5;

Qy 1 ETFFPKYLHYDEETSHQLLCKDKCPGTYLKQCTAKWTKVACAPCPDHYTDSWHTSDECL 60  
Db 1 ETFFPKYLHYDEETSHQLLCKDKCPGTYLKQCTAKWTKVACAPCPDHYTDSWHTSDECL 60  
Qy 61 YCSPVCKELQYVQECNTRNHRVCECKEGRYLEIEFCLKLRHSCPPGFGVQAGTPERNTV 120  
Db 61 YCSPVCKELQYVQECNTRNHRVCECKEGRYLEIEFCLKLRHSCPPGFGVQAGTPERNTV 120  
Qy 121 CKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTOKVDKTHTC 180  
Db 121 CKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSESTOK 173  
Qy 181 PPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 240  
Db 174 --CGIDVTLCCEAFAFRFAVPTK-----FTPNWLSVLVD---AESVE--- 217

Qy 241 AKTKPREQYNSTYRVSV 259  
Db 218 -RIKROHSSQEOTFOLLKL 235

## RESULT 3

Q08712 PRELIMINARY; PRT; 401 AA.  
AC Q08712; 070202;  
DT 01-JUL-1997 (Tremblrel. 04, Created)  
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)  
DE (OCIF).  
GN TNFRSF11B OR OPG.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]

SEQUENCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE-KIDNEY;  
 RX MEDLINE=97262071; PubMed=9108485;  
 RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,  
 RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,  
 RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,  
 RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,  
 RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,  
 RA Suggs S., Boyle W.J.,  
 RA "Osteoprotegerin: a novel secreted protein involved in the regulation  
 of bone density";  
 RL Cell 89:309-319(1997).  
 RN [2]  
 SEQUENCE FROM N.A.  
 RC STRAIN=129/OLA, AND NIH SWISS;  
 RX MEDLINE=98382527; PubMed=9714833;  
 RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,  
 RA Higashio K.;  
 RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)  
 gene and its expression in embryogenesis";  
 RL Gene 215:339-343(1998).  
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
 OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY  
 SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
 OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
 STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
 CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).  
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,  
 BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND  
 PLACENTA. NOT DETECTED IN SPLEEN.  
 CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT  
 DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY  
 15 TO DAY 17.  
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; U94331; AAB53708.1; -;  
 DR EMBL; AB013898; BAA28269.1; -;  
 DR EMBL; AB013903; BAA33388.1; -;  
 DR EMBL; AB013899; BAA33388.1; JOINED.  
 DR EMBL; AB013900; BAA33388.1; JOINED.  
 DR EMBL; AB013901; BAA33388.1; JOINED.  
 DR EMBL; AB013902; BAA33388.1; JOINED.  
 DR HSSP; P25942; 1CDF.  
 DR MGD; MGI:109587; Opg.  
 DR INTERPRO: IPR000488; -;  
 DR INTERPRO: IPR001368; -;  
 DR PFAM; PF00020; TNFR\_C6; 3.  
 DR PROSITE; P500652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; P50017; DEATH\_DOMAIN; 1.  
 DR PROSITE; P50050; TNFR\_NGFR\_2; 2.  
 DR PRODOM; PD000771; -; 1.  
 KW Glycoprotein; Repeat; Cytokine; Signal.  
 FT SIGNAL  
 FT CHAIN 1 21  
 FT DOMAIN 22 401 OSTEOPROTEGERIN.  
 FT REPEAT 23 201 4 X TNFR-CYS.  
 FT REPEAT 23 63 TNFR-CYS 1.  
 FT REPEAT 64 106 TNFR-CYS 2.  
 FT REPEAT 107 143 TNFR-CYS 3.  
 FT REPEAT 144 201 TNFR-CYS 4.  
 FT DOMAIN 306 365 DEATH DOMAIN.  
 FT DISULFID 41 54 BY SIMILARITY.  
 FT DISULFID 44 62 BY SIMILARITY.  
 FT DISULFID 65 80 BY SIMILARITY.  
 FT DISULFID 83 97 BY SIMILARITY.  
 FT DISULFID 87 105 BY SIMILARITY.  
 FT DISULFID 118 142 BY SIMILARITY.  
 FT DISULFID 145 160 BY SIMILARITY.  
 FT CARBOHYD 98 98 POTENTIAL.  
 FT CARBOHYD 165 165 POTENTIAL.  
 FT CARBOHYD 178 178 POTENTIAL.  
 FT CARBOHYD 289 289 POTENTIAL.  
 FT VARIANT 138 138  
 FT R -> P (IN STRAINS 129/OLA AND NIH  
 FT SWISS).

FT VARIANT 161 161 I -> R (IN STRAINS 129/OLA AND NIH  
 FT SWISS).  
 FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH  
 FT SWISS).  
 FT VARIANT 288 288 S -> A (IN STRAINS 129/OLA AND NIH  
 FT SWISS).  
 FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH  
 FT SWISS).  
 SQ SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;  
 Query Match 39.9%; Score 893.5; DB 11; Length 401;  
 Best Local Similarity 47.8%; Pred. No. 1.1e-70;  
 Matches 191; Conservative 35; Mismatches 107; Indels 67; Gaps 9;  
 QY 1 ETTPPKYLHDEETSHQLCDKCPGPGTYLKHCHTAKWKTVCAPCPDHVYTDSWHTSDECL 60  
 DB 22 ETLPKYLHVDPTGHQLCDKCAPGTYLKHCHTAKWKTVCAPCPDHVYTDSWHTSDECL 81  
 QY 61 YCSPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHSRCPGPGVQAGTPERTV 120  
 DB 82 YCSPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHSRCPGPGVQAGTPERTV 141  
 QY 121 CKRCPDFFSNETSSKAPCRKHTNCVFGLLLTQKGNATHDNCISGNESTQKVDKTHTC 180  
 DB 142 CKKCPDFFSNETSSKAPCRKHTNCVFGLLLTQKGNATHDNCISGNESTQKVDKTHTC 194  
 QY 181 PCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEHN 240  
 DB 195 --CGIDVTLCCEAFFFAVPTK-----IIPNMLSVLVD---SLPQKVN--AESVE--- 238  
 QY 241 AKTKPREQNSYRVVSVLTVLHQLDNLNGKEY-----CKKVS----- 278  
 DB 239 -RKRHRSSQEQFQLLKWKHQNRQEMVYKIIQDIDLCSSVQRHLGHSNLTTEQLLA 297  
 QY 279 -----NKALPAPIEKTISKAKGPQPVQTYL-----PPSRDELTKNQVSLTCLVKGFT 327  
 DB 298 LMESLPKGIKISPEIETRTKTSKESQQLKLSLRKNGDQDTLKLGLVYALKHLKTSHE 357  
 QY 328 PSDIADVWESNGQPNENYKTPPVLDSDGSGFFLYSKLTVD 367  
 DB 358 PKTVT-----HSLRKTMRFLSHFTMYRLYQKLFLE 387  
 RESULT 4  
 O08727 PRELIMINARY; PRT; 401 AA.  
 ID O08727  
 AC O08727;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)  
 DE (OCIF).  
 GN TNFRSF11B OR OPG.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_taxid=10116;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=INTESTINE;  
 RX MEDLINE=97262071; PubMed=9108485;  
 RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,  
 RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,  
 RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,  
 RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,  
 RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,  
 RA Suggs S., Boyle W.J.,  
 RA "Osteoprotegerin: a novel secreted protein involved in the regulation  
 of bone density";  
 RL Cell 89:309-319(1997).  
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
 OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY

CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
 CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
 CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
 CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY  
 CC SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; U94330; AAB53707.1; -.  
 DR HSSP; P25942; ICDP.  
 DR INTERPRO; IPR001368; -.  
 DR PFAM; PF00020; TNFR\_C6; 4.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; PS00050; TNFR\_NGFR\_2; 2.  
 DR PRODOM; PD000771; -. 1.  
 KW Glycoprotein; Repeat; Cytokine; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 401 BY SIMILARITY.  
 FT DOMAIN 23 201 OSTEOPROTEGERIN.  
 FT REPEAT 23 63 4 X TNFR-CYS.  
 FT REPEAT 64 106 TNFR-CYS 1.  
 FT REPEAT 107 143 TNFR-CYS 2.  
 FT REPEAT 144 201 TNFR-CYS 3.  
 FT REPEAT 202 263 TNFR-CYS 4.  
 FT DOMAIN 306 365 DEATH DOMAIN.  
 FT DISULFID 41 54 BY SIMILARITY.  
 FT DISULFID 44 62 BY SIMILARITY.  
 FT DISULFID 65 80 BY SIMILARITY.  
 FT DISULFID 83 97 BY SIMILARITY.  
 FT DISULFID 87 105 BY SIMILARITY.  
 FT DISULFID 118 142 BY SIMILARITY.  
 FT DISULFID 145 160 BY SIMILARITY.  
 FT CARBOHYD 98 98 POTENTIAL.  
 FT CARBOHYD 165 165 POTENTIAL.  
 FT CARBOHYD 178 178 POTENTIAL.  
 FT CARBOHYD 289 289 POTENTIAL.  
 SQ SEQUENCE 401 AA; 46192 MW; FEC6A31F1D4E573A CRC64;

Query Match 39.2%; Score 877; DB 11; Length 401;  
 Best Local Similarity 63.7%; Pred. No. 3.2e-69;  
 Matches 165; Conservative 18; Mismatches 52; Indels 24; Gaps 5;

Qy 1 ETFPPKYLHYDEEESHQLLCKPCPTYLKQHCYAKWTKVACPDHYHTSWHTSDCL 60  
 Db 22 ETFPPKYLHYDEEESHQLLCKPCPTYLKQHCYAKWTKVACPDHYHTSWHTSDCL 81  
 Qy 61 YCSPVKELQVQKQCNTHNRVCKEGRYLEIEFCLKHRSPPGFGVQAGTPERNIV 120  
 Db 82 YCSPVKELQVQKQCNTHNRVCKEGRYLEIEFCLKHRSPPGFGVQAGTPERNIV 141  
 Qy 121 CKRCPDGFSSNETSKAPCRKHTNCSVFLGLLTQKGNATHDNCISGNSSESTQKVDKTHC 180  
 Db 142 CKRCPDGFSSNETSKAPCRKHTNCSVFLGLLTQKGNATHDNCISGNSSESTQKVDKTHC 194  
 Qy 181 PPCPAPELLGSPVLEFPKPKDILMISRTVEVTCVVDVSHEDPEVKFNVDGVEVHN 240  
 Db 195 --CGIDVTLCGEAFRAVPTK-----IIPNLSVLVD---SLPGTKVN--AESVE--- 238  
 Qy 241 AKTKPREQYNSTYRVVSV 259  
 Db 239 -RIKRRHSSQEQTFQLKL 256

RESULT 5  
 Q9RIA4 PRELIMINARY; PRT; 437 AA.  
 AC Q9RIA4:  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DT GAMMAL HEAVY CHAIN OF MAB7 (FRAGMENT).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;  
 RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal  
 RT antibody (Mab 7, its light and heavy chains) and construction of a  
 RT single chain antibody (scFV).";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF152372; AAD40243.1; -.  
 DR HSSP; P01842; 7FAB.  
 DR INTERPRO; IPR003006; -.  
 DR PFAM; PF00047; Ig\_4.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
 FT NON\_TER 1 437  
 FT NON\_TER 437 437  
 SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;  
 Query Match 36.9%; Score 827; DB 11; Length 437;  
 Best Local Similarity 60.9%; Pred. No. 8.8e-65;  
 Matches 145; Conservative 44; Mismatches 39; Indels 10; Gaps 3;  
 Qy 170 STQKVDKTHT-----CPP--CPAPELLGSPVLEFPKPKDILMISRTVEVTCVVDVSH 222  
 Db 202 SSTKVDDKKIVPRDCGCKPCICTVPEV---SSVFIEFPKPKDILMISRTVEVTCVVDVSH 258  
 Qy 223 EDPEVKFNWYDGVVHNKTKPREQYNSTYRVVSVLTVLHODWLNKGEYCKYKSNKAL 282  
 Db 259 DDPEVQFSWFDVDDVEHTAQTPREEQNFSTRSSELPIMHQDWLNKGEYCKYKSNKAL 318  
 Qy 283 PAPIEKTISKAKGPPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGOPE 342  
 Db 319 PAPIEKTISKAKGPPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGOPE 378  
 Qy 343 NNYKTPPVLDSDGFFLYSKLTVDKSRWQQGNFVSCSVMHAEALHNHYTKSLSPG 400  
 Db 379 ENYKNTQPIMDTGSYFYVSKLVQKSNWEAGNTFTCSVLHDEGLHNHTKNSLSPG 436  
 RESULT 6  
 O95407 PRELIMINARY; PRT; 300 AA.  
 AC O95407:  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).  
 GN DCR3 OR TR6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel P.C.,  
 RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,  
 RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,  
 RA Goddard A.D., Botstein D., Ashkenazi A.;  
 RT "Genomic amplification of a decoy receptor for Fas ligand in lung and  
 RT colon cancer.";  
 RL Nature 396:699-703(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BLOOD;  
 RX MEDLINE=99253915; PubMed=10318773;  
 RA Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;  
 RT "A newly identified member of tumor necrosis factor receptor  
 RT superfamily (TR6) suppresses LIGHT-mediated apoptosis.";  
 RL J. Biol. Chem. 274:13733-13736(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PANCREAS;



```

RX MEDLINE-20122600; PubMed-10655513;
RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT "Overexpression of M8/DCR3 in human gastrointestinal tract tumors
RT independent of gene amplification and its location in a four-gene
RT cluster.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
RL EMBL; AF104419; AAD03056.1; -.
DR EMBL; AF134240; AAD29688.1; -.
DR EMBL; AF217796; AAF35244.1; -.
DR EMBL; AF217793; AAF33685.1; -.
DR EMBL; AF217794; AAF33686.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_NGFR_1; UNKNOWN_1..
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1..
DR PROSITE; PS01186; EGF_2; UNKNOWN_1..
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match
Best Local Similarity 19.3%; Score 433; DB 4; Length 300;
Matches 69; Conservative 29; Mismatches 63; Indels 0; Gaps 0;

QY 5 PKYLHDEETSHQLLCKDPPGTYLKQHTAKWKTCVACPDHYHDTDSWHTSDECLYCSP 64
DB 34 PTYPWRDAETGERLVAQCQPGTFVQPCRRDSPTTCGCPGPRHYTFQWYLERCRYCNV 93
QY 65 VKLEIVYKQECNRTNHRVCECKEGRYL--ETFECLKHRSCPPGFGVQAGTPERTNTVCKRC 124
DB 94 LCGEREERACHATNRCRGTGFAGHAGFCLEHACPPGAGVIAPGTPSQNTQCQC 153
QY 125 PDGFFSNETSAPCRKHTNCSVFGLLLTQKGNATHDNICS 165
DB 154 PPGTFSSASSSSQCPHRNCTALGLALNVPSSSHDTLCT 194

RESULT 7
ID Q9PUS0 PRELIMINARY; PRT; 302 AA.
AC Q9PUS0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE DECOY RECEPTOR.
OS Salvelinus fontinalis (Brook trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN [1]
RP SEQUENCE FROM N.A.
RA Bohe J., Goetz F.W.;
RT "A tumor necrosis factor receptor homolog is up-regulated in the brook
RT trout (Salvelinus fontinalis) ovary at the completion of ovulation.";
RL Biol. Reprod. 0:0-0(1999).
DR EMBL; AF156738; AAD56428.1; -.
DR HSSP; P19438; 1EXT.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_NGFR_1; UNKNOWN_1..
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1..
DR PROSITE; PS01186; EGF_2; UNKNOWN_1..
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match
18.1%; Score 405.5; DB 13; Length 302;

```

```

Best Local Similarity 45.2%; Pred. No. 6.7e-28;
Matches 71; Conservative 30; Mismatches 53; Indels 3; Gaps 2;

QY 11 DEETSHQLLCKDPPGTYLKQHTAKWKTCVACPDHYHDTDSWHTSDECLYCSPVKELQ 70
DB 27 DRYSGLSIVCDRCPPGTYLRAPCSAMRKSDCAECPNAGYTFEWNHITSKLCRCS-MCAENQ 85
QY 71 YVKECNRTNHRVCECKEGRYL--ETFECLKHRSCPPGFGVQAGTPERTNTVCKRCPDGF 128
DB 86 VVKECSPSNCECKEGRYFNKKYEACIKHKECPGCGYANTTGTGPHQTECVQCOAGF 145
QY 129 FSNETSAPCRKHTNCSVFGLLLTQKGNATHDNICS 165
DB 146 YSEVSSAKATCLAQSNCKVGLRVLLKGQDWHNTLCA 182

RESULT 8
Q62327 PRELIMINARY; PRT; 459 AA.
ID Q62327;
AC Q62327;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
RT gene.";
RL Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -.
DR HSSP; P19438; 1NCF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_NGFR_1; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
FT NON_TER 1
FT VARIANT 87 S -> T.
FT VARIANT 93 T -> I.
FT VARIANT 268 F -> I.
FT VARIANT 345 S -> F.
FT VARIANT 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match
15.4%; Score 345; DB 11; Length 459;
Best Local Similarity 25.2%; Pred. No. 2.3e-22;
Matches 100; Conservative 53; Mismatches 164; Indels 80; Gaps 12;

QY 9 HYDEETSHQLLCKDPPGTYLKQHTAKWKTCVACPDHYHDTDSWHTSDECLYCSPVCKE 68
DB 31 YDORKA--QMCACKCPGQGVKHFNCNTSDTVCADEASMTQVWNFRCLSCSSCSST 88
QY 69 LQYVKQECNRTNHRVCECKEGRYLEIEF-----CLKHRSCPPGFGVQAGTPERTNTVC 121
DB 89 DQVETRACCTQQNRVCACAGRYCALKTHSGSCRCQMRSLKCGCGFGCVASSRAENGVL 148
QY 122 KRCPDGFFSNETSAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQKV----- 174
DB 149 KACAPGTFSDTSDVCRPHRICSLAI----PGNASTDAVCAPESPTLSAIPRTLYVS 204
QY 175 --DKHTTCP-----PCPAPELL-----GGPSV-----FLEPPK 200
DB 205 QPEPTRSQPLDQEPGFSQTPSILTSLGSTPIIEOSTKGGISLPIGLIVGVTSLGLMLGL 264
QY 201 PKDTLMSRPEVTVCVVVDVSHEDPEVKFNWYVDGVGVHNAKTPREEQYNSTYRVYSVL 260

```

Db 265 VNCFILVQRKKKPSCLQDA--KVPHVDEKSDQAVGL-----EQQH-----LL 306  
QY 261 TVLHQDWLNGKCKKYSNKAIPA---PIEKTISKAGQPREQVYTLPSRDELTKNOV 317  
Db 307 TTPAPSSSSSLESASAGDRAPPGGHPQARVMAEAGSQEASRRSSRSDSHGSHGTHV 366  
QY 318 SLTCLVKGYFSPDIAVWESN-----GQPNNTKTP 349  
Db 367 NVTCIIVNCSSSDHSSQCSQASATVGDPAKPSASP 403  
RESULT 9  
Q16042 PRELIMINARY; PRT; 439 AA.  
AC Q16042;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)  
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE TUMOR NECROSIS FACTOR RECEPTOR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91370690; PubMed=1966549;  
RA Dembic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R.,  
RA Brockhaus M., Lesslauer W.;  
RT "Two human TNF receptors have similar extracellular, but distinct  
RT intracellular, domain sequences."  
RL Cytochrome 2:231-237(1990).  
DR EMBL; S63368; AAB19824.1; .  
DR HSSP; P25942; 1CDF.  
DR INTERPRO; IPR001368; .  
DR PFAM; PF00020; TNFR\_C6; 4.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.  
DR PROSITE; PS00050; TNFR\_NGFR\_2; 3.  
DR PRODOM; PD000771; .; 1.  
SQ SEQUENCE 439 AA; 46090 MW; FEBCBEB329CC67FF6 CRC64;

Query Match 15.3%; Score 343.5; DB 4; Length 439;  
Best Local Similarity 26.8%; Pred. No. 3e-22;  
Matches 107; Conservative 44; Mismatches 163; Indels 85; Gaps 15;  
QY 9 HYDETSHQLLCKPCPGTYLKQHTCAKWKVACPCPDHYTDSWHTSDECLY 68  
Db 23 YDQ--TAQMCCKSCSPGQHAKVFCTKTSYVCDSCEDSTYQLWNVPECLSCGSCSS 80  
QY 69 LQYKQECNRTHNRVCECKEGRYLEI-----EFLKHSRCPGPGVVOAGTPERTVCK 122  
Db 81 DQVETQACTREQNRICTRPGNYCALSKQEGRCRLCAPLRCRPGFGVAPGTETSDVYCK 140  
QY 123 RCPDGFNSETSSKAPCRKHTNCSVFGLLHKGKGNATHNICSGNES--TKVDKTRCP 181  
Db 141 PCAPGTFSNTSTSDICRPHQCNVVAI-----PGNASMDVACTSTPSRMAPAGVHLPO 196  
QY 182 PC-----PAPELLGPGSV-FLFPPKPKDT-----LMISRT-----E 212  
Db 197 PVSTRSQHTQPTPEFTAPSTSFLLPMGPSPPAEGSTGDFALPVGLIVGVTLGLLIGV 256  
QY 213 VTCVVVDVSHEDP-----EVFNKYVDQVEVHNAK--TKPREQYNSTYRVVSVLTVLHQD 266  
Db 257 VNCVIMTVQKKKPLCLQREAK-----VPHLPADKARGTQGPQQH-----LLITAPSS 304  
QY 267 WLNGKEYCKYSNKAIPAEIKTISKAGQPREQVYTLPSRDELTK-----N 315  
Db 305 SSSLESSASALDRAP-----TRNQFAPGVEASGAGEARASTGSSDSFPGGHGT 355  
QY 316 QVSLTCLVKGYFSPDIAVWESN-----GQPNNTKTP 349  
Db 356 QVNVTCIIVNCSSSDHSSQCSQASSTMGDTSDSPSPSP 394

RESULT 10  
O88734 PRELIMINARY; PRT; 482 AA.  
ID O88734;  
AC O88734;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE P80 TNF-ALPHA RECEPTOR.  
GN TNFR2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hurle B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;  
RT "The Mouse Tumor Necrosis Factor Receptor Gene:Genomic Structure and  
RT Characterization of the two Transcripts."  
RL Genomics 0:0-0(0).  
DR EMBL; Y14619; CAA74969.1; .  
DR EMBL; Y14620; CAA74969.1; JOINED.  
DR EMBL; Y14621; CAA74969.1; JOINED.  
DR EMBL; Y14622; CAA74969.1; JOINED.  
DR EMBL; Y14623; CAA74969.1; JOINED.  
DR EMBL; Y14679; CAA74969.1; JOINED.  
DR HSSP; P19438; 1NCF.  
DR INTERPRO; IPR001368; .  
DR PFAM; PF00020; TNFR\_C6; 4.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.  
DR PROSITE; PS00050; TNFR\_NGFR\_2; 3.  
DR PRODOM; PD000771; .; 1.  
SQ SEQUENCE 482 AA; 51106 MW; F6C15046B48FF83C CRC64;  
Query Match 15.1%; Score 338; DB 11; Length 482;  
Best Local Similarity 25.1%; Pred. No. 1e-21;  
Matches 102; Conservative 50; Mismatches 164; Indels 90; Gaps 13;  
QY 9 HYDETSHQLLCKPCPGTYLKQHTCAKWKVACPCPDHYTDSWHTSDECLY 61  
Db 46 YDKRA--QMCCKACPPQYVHFCNKTSYVCDSDIVCADCEASMTQVWNOFRTCLS 103  
QY 62 CSPVCKELQYKQECNRTHNRVCECKEGRYLEIEF-----CLKHSRCPGPGVVOAGT 114  
Db 104 CSSSCSTDQVETRACTKQONRVCAACEAGRYCALKTHSGRCQCMRLSKGPGFGVASSRA 163  
QY 115 PERNTVCRCPDGFNSETSSKAPCRKHTNCSVFGLLHKGKGNATHNICSGNESSTQKV 174  
Db 164 PNGNVLCACAPGTFSDTSTSDVCRPHRICSLAI-----PGNASTDVCAPESPTLSAI 219  
QY 175 -----DKTHTCP-----PCPAPELL-----GGPSV----- 194  
Db 220 PRTLYVSQPEFTRSQPLDQEPGPGSQTPSILTSLSGTPPIEQSTKGISLPLGLIVGVTS 279  
QY 195 -FLFPPKPKDTLMISRTPEVTCVVVD--VSHEDPEVKFNWYVDGVEVHNAKTKPREQYN 251  
Db 280 GLLMGLVNCVILVORKKPKSCLOQDAKVPVH--PDEKSQDAGVLEQQHLLITAPSSSSS 338  
QY 252 STYRVSVSLTVLHQDWLNGKEYCKYSNKAIPA---PIEKTISKAGQPREQVYTLPPS 308  
Db 339 SL-----ESSASAGDRAPPGGHPQARVMAEAGSQEASRRSSRSDS 380  
QY 309 RDELTKNOVSLTCLVKGYFSPDIAVWESN-----GQPNNTKTP 349  
Db 381 SHGSHGTHVNVTCIIVNCSSSDHSSQCSQASATVGDPAKPSASP 426  
RESULT 11  
O75509 PRELIMINARY; PRT; 655 AA.  
ID O75509;  
AC O75509;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE TNFR-RELATED DEATH RECEPTOR-6 (DJ181J13.1) (DR6 (TNFR-RELATED DEATH  
DE RECEPTOR-6)).  
GN DR6 OR DJ181J13.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fan G., Bauer J.H., Haridas V., Wang S., Liu D., Ni J., Yu G.,  
RA Vincenz C., Aggarwal B.B., Dixit V.M.;  
RT "Identification and functional characterization of DR6, a novel death  
RT domain-containing TNF receptor";  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Parker A.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF068868; AAC34583.1; -;  
DR EMBL; AL096801; CAB75692.1; -;  
DR HSSP; P07174; INGR.  
DR INTERPRO; IPR000488; -;  
DR INTERPRO; IPR001368; -;  
DR PFAM; PF00020; TNFR\_c6; 4.  
DR PFAM; PF00531; death\_1.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
DR PROSITE; PS0017; DEATH\_DOMAIN; 1.  
DR PROSITE; PS00050; TNFR\_NGFR\_2; 1.  
DR PRODOM; PD000771; -; 1.  
KW Receptor.  
SQ SEQUENCE 655 AA; 71844 MW; 48939391C4852A33 CRC64;

Query Match 13.3%; Score 297; DB 4; Length 655;  
Best Local Similarity 28.3%; Pred. No. 5.9e-18;  
Matches 87; Conservative 44; Mismatches 134; Indels 42; Gaps 11;

QY 7 YLHYDEETSHQLCDKCPPTYLKQHCTAKWTVACAPDPHYTDSWHTSDECLYCSVPC 66  
Db 54 YRVHVRATGQVLTCDCPCAGTYVSEHCTNLSRVCSCPVGTTRHENGIEKCHDCSQPC 113  
QY 67 KELQYVQECNRTHNRVCECKEGRYLEIEPCLKHRSPPGFGVVGQAGTPERNTVCKRCPD 126  
Db 114 PWFMIKLPAAALTDRECTCPGCFQSNATCAPHTVCPVGWGVKKGKTETEDVRCKOCAR 173  
QY 127 GFESNETSSKAPCKRKHNCVSFGLLLTKQGNATHDNICS---GNSESTQKVDKTHTCPPC 183  
Db 174 GTESDVPSSVMKCAVTDCLSQLNLVWIKPGTKETDNNVCGTLPFSFSSTS----- 222  
QY 184 PAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHED-----PEVKFNWYVDG 235  
Db 223 PSP-----GTAIF---PRP-EHMETHEVPSTTYYPKGMNSTESNASSVRPKV-LSSIQEG 273  
QY 236 VEVHNATKPREQYNSTYRWVSLTVLHODWLNGLNGKEYCKVSNKALPAPI-----EKTIS 291  
Db 274 TVPDNTSSARGKEDVNTK---LPNLQVNVHQ---QGPHRHIL--KLLPSMEATGGEKSS 326  
QY 292 KAKGQPR 298  
Db 327 PIKGPKR 333

RESULT 12  
Q90544  
ID Q90544 PRELIMINARY; PRT; 684 AA.  
AC Q90544;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.  
OS Ginglymostoma cirratum (Nurse shark).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;  
OC Ginglymostomatidae; Ginglymostoma.  
OX NCBI\_TaxID=7801;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RX MEDLINE=951831140; PubMed=7877689;  
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,  
RA Flajnik M.F.;  
RT "A new antigen receptor gene family that undergoes rearrangement and  
RT extensive somatic diversification in sharks";  
RL Nature 374:168-173(1995).  
DR EMBL; U18701; ABA48195.1; -;  
DR HSSP; P01857; 1FC1  
DR INTERPRO; IPR003006; -;  
DR PFAM; PF00047; I9; 6.  
DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_3.  
KW Signal.  
FT SIGNAL 1 18 POTENTIAL.  
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.  
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;

Query Match 12.5%; Score 280.5; DB 13; Length 684;  
Best Local Similarity 26.5%; Pred. No. 1.7e-16;  
Matches 103; Conservative 53; Mismatches 148; Indels 85; Gaps 18;

QY 51 DSWHTSDECLYCS-----PVCLEQYVQECNRTHNRVCECKEGRYLEIEPCLKHRS 102  
Db 315 EEWQSGVE-YTCSAKQDQSSSTPVVVKTRKARVEPTKPLRL----- 354  
QY 103 CPFGFVGWQAGTPERNTVCKRCPDGFESNETS---SKAPCKRKHNCVSFGLLLTQ----- 154  
Db 355 LPSPPEIQSTSSATLTLCLIR---GFYPDKVSVSWQKDDVSVSANVTNFTALEQDLTF 411  
QY 155 -----KGNATHDNICGN---SEST-----QKVDKTHTCPPCAPPELLGGP 192  
Db 412 TRSLNLNTAVENKSGAKY--TCTASHPPSQSTVYKRVIRNQVD-----CROPDI-- 458  
QY 193 SVLEFPKPKDTLMISRTPEVTCVVVDVSHEDPE-VKFNWYVDGVEVHNATKPREQYN 251  
Db 459 SVSLLKP-PPEEIWTQQTATIVCEIV---YSDLENIKVFWQVNGVERKKGVTQNPWSG 514  
QY 252 STYRVSVLTVLHODWLNGLNGKEYCKVSNKALPAPIETISKAK-GPREQVYTLPPSRD 310  
Db 515 SKSTIVSKLVMASEWSDSGTEYVCLVEDSELPTVKASIRKANYSQMHPPKVYLLHPSTD 574  
QY 311 EL-TKNOVSLTCLVKGFYPSDIAVWEWSNGO-PENNYKTTTPVLDSDSGSFYLSKLTVDK 368  
Db 575 EIDTENSATLMCLATNHPAEIYGVWMANDTLSDSGYRTQVDSKSGSFVTDRLRLTA 634  
QY 369 SRWQQGVNFCVSMHEALHN---HYTQKS 394  
Db 635 AEWNSDTTYSCLVGHPSLNRLDIRSTNKS 663

RESULT 13  
Q9UP60  
ID Q9UP60 PRELIMINARY; PRT; 384 AA.  
AC Q9UP60;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE SNC73 PROTEIN.  
GN SNC73.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zheng S., Cao J., Cao W., Cai X., Geng L.;



```

FT DISULFID 50 68 BY SIMILARITY.
FT DISULFID 71 86 BY SIMILARITY.
FT DISULFID 92 112 BY SIMILARITY.
FT DISULFID 114 124 BY SIMILARITY.
FT DISULFID 126 133 BY SIMILARITY.
FT DISULFID 127 131 BY SIMILARITY.
FT DISULFID 154 169 BY SIMILARITY.
FT DISULFID 175 194 BY SIMILARITY.
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 616 AA; 66033 MW; E3DE9A7A08196F81 CRC64;

Query Match 11.5%; Score 258.5; DB 4; Length 616;
Best Local Similarity 34.5%; Pred. No. 1.3e-14;
Matches 58; Conservative 21; Mismatches 82; Indels 7; Gaps 4;

QY 20 CDKCPGTYLKQCHTAKWTVACPDHYTDSWHTSDECLYCPVC---KELOYVKQEC 76
Db 47 CNKCEFGKYMSSKCTTTSDSVCLPCGPFDEYLDWSNEEDKCL-LHKVCDTGKALVAVVAG- 104

QY 77 NRTHNRVCECKEGRY--LEIEFCLKHRSCPPGFGVVQAGTPERTVCKRCPCDGFESNETS 134
Db 105 NSTTPRRCACTAGYHWSQDCECCRRNTECAPGLGAQHPLQLNKDTVCKPCLAGYFSDAFS 164

QY 135 SKAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQKVDKTHTCPP 182
Db 165 STDKCRPWNTCTFLGKRVEHHGTEKSDAVCSSSLPARKPPNEPHVYLP 212

```

Search completed: March 1, 2001, 09:20:07  
Job time: 405 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:44 ; Search time 40.97 seconds  
(without alignments)  
315.295 Million cell updates/sec

Title: US-09-389-782A-5

Perfect score: 2240

Sequence: 1 ETTPPKYLHYDEETSHQLC.....VMHEALHNHYTQKSLSPG 400

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1234	55.1	330	1 GCL_HUMAN	P01857 homo sapien
2	1148.5	51.3	326	1 GC2_HUMAN	P01859 homo sapien
3	1136	50.7	327	1 GC4_HUMAN	P01861 homo sapien
4	1133.5	50.6	290	1 GC3_HUMAN	P01860 homo sapien
5	923	41.2	323	1 GC_RABIT	P01870 oryctolagus
6	904.5	40.4	329	1 GC2_CAVPO	P01862 cavia porce
7	847.5	37.8	329	1 GC3_MOUSE	P22436 mus musculu
8	843	37.6	333	1 GCB_RAT	P03987 rattus norv
9	841.5	37.6	398	1 GC3M_MOUSE	P01868 mus musculu
10	820	36.6	324	1 GCL_MOUSE	P20759 rattus norv
11	820	36.6	326	1 GCL_RAT	P01869 mus musculu
12	820	36.6	393	1 GC1M_MOUSE	P01863 mus musculu
13	812.5	36.3	330	1 GCRA_MOUSE	P01865 mus musculu
14	812.5	36.3	399	1 GCAM_MOUSE	P20762 rattus norv
15	804.5	35.9	329	1 GCC_RAT	P01864 mus musculu
16	801.5	35.8	335	1 GCAB_MOUSE	P20760 rattus norv
17	791	35.3	322	1 GCA_RAT	P01866 mus musculu
18	777.5	34.7	336	1 GCB_MOUSE	P01867 mus musculu
19	777.5	34.7	405	1 GCBM_MOUSE	P04220 homo sapien
20	365.5	16.3	391	1 MUCB_HUMAN	P06336 mus musculu
21	361	16.1	421	1 EPC_MOUSE	P01871 homo sapien
22	360	16.1	454	1 MUC_HUMAN	P01872 mus musculu
23	354.5	15.8	455	1 MUC_MOUSE	P01854 mus sapien
24	353	15.8	428	1 EPC_HUMAN	P01855 rattus norv
25	353	15.8	429	1 EPC_RAT	P01873 mus musculu
26	349.5	15.6	476	1 MUCM_MOUSE	P20333 homo sapien
27	345.5	15.4	461	1 TNR2_HUMAN	P03988 oryctolagus
28	340	15.2	458	1 MUC_RABIT	P25119 mus musculu
29	340	15.2	474	1 TNR2_MOUSE	P01874 canis fami
30	338.5	15.1	450	1 MUC_CANFA	P04221 oryctolagus
31	335	15.0	479	1 MUCM_RABIT	P06337 mesocricetu
32	333.5	14.9	454	1 MUC_MESAU	P20768 suncus muri
33	332.5	14.8	457	1 MUC_SUNNU	

#### ALIGNMENTS

RESULT 1

ID	GCL_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG GAMMA-1 CHAIN C REGION.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF I-135 (MYELOMA PROTEIN EU).			
RP	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RL	Biochemistry 9:3161-3170(1970).			
RN	[3]			
RP	SEQUENCE OF 136-329 (EU).			
RP	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RL	Biochemistry 9:3171-3181(1970).			
RN	[4]			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RP	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein NIE) III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RN	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RP	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[6]			
RP	DISULFIDE BONDS			
RP	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X. Intrachain disulfide bonds."			

P23085 heterodontu  
P23087 heterodontu  
P01879 oryctolagus  
P23088 heterodontu  
P23086 heterodontu  
P27512 mus musculu  
P01875 gallus gall  
P23084 heterodontu  
P01876 homo sapien  
P20758 gorilla gor  
P01877 homo sapien  
P25942 homo sapien

34 315.5 14.1 438 1 HVC2\_HETFR  
35 299 13.3 438 1 HVCS\_HETFR  
36 298 13.3 299 1 ALC\_RABIT  
37 295 13.2 461 1 HVCN\_HETFR  
38 288 12.9 393 1 HVC3\_HETFR  
39 287.5 12.8 289 1 CD40\_MOUSE  
40 282.5 12.6 446 1 MUC\_CHICK  
41 279.5 12.5 370 1 HVC1\_HETFR  
42 277.5 12.4 353 1 ALC1\_HUMAN  
43 274.5 12.3 353 1 ALC1\_GORGO  
44 273.5 12.2 340 1 ALC2\_HUMAN  
45 264 11.8 277 1 CD40\_HUMAN

RL Biochemistry 9:3188-3196(1970).  
RN [7]  
RP DISULFIDE BONDS.  
RX MEDLINE=77070267; PubMed=1002129;  
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
RT "Rule of antibody structure. The primary structure of a monoclonal  
RT IgG1 immunoglobulin (myeloma protein N1e), I: Purification and  
RT characterization of the protein, the L- and H-chains, the  
RT cyanogen bromide cleavage products, and the disulfide bridges";  
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RX MEDLINE=81208100; PubMed=7236608;  
RA Deisenhofer J.;  
RT "Crystallographic refinement and atomic models of a human Fc fragment  
RT aureus at 2.9- and 2.8-A resolution.";  
RL Biochemistry 20:2361-2370(1981).  
CC -!- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE  
CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)  
CC MARKER & THE GIM (NON-1) MARKERS.  
CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
CC 35,116,198,269 & 272.  
CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
CC 268-272.  
CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
CC RESIDUES 198,267&272.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J00228; AAC82527.1; ALT\_INIT.  
DR PIR; A02146; GHU.  
DR PDB; 1FC1; 15-JUL-92.  
DR PDB; 1FC2; 15-JUL-92.  
DR MIM; 147100; -.  
DR INTERPRO; IPR000495; -.  
DR INTERPRO; IPR003006; -.  
DR PFAM; PF00047; 1g; 3.  
DR PROSITE; PS00290; IG\_MHC; 2.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
KW 3D-structure.  
FT NON\_TER 1 1  
FT DOMAIN 1 98 CH1.  
FT DOMAIN 99 110 HINGE.  
FT DOMAIN 111 223 CH2.  
FT DOMAIN 224 330 CH3.  
FT DISULFID 27 83  
FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).  
FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).  
FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).  
FT DISULFID 144 204  
FT DISULFID 250 308  
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .).  
FT VARIANT 97 97 /FTid=VAR\_003886.  
FT VARIANT 239 239 D -> E (IN GIM(NON-1) MARKER).  
FT VARIANT 241 241 /FTid=VAR\_003887.  
FT MOD\_RES 330 330 L -> M (IN GIM(NON-1) MARKER).  
FT STRAND 123 126 /FTid=VAR\_003888.  
FT HELIX 130 134 REMOVED POST-TRANSLATIONALLY.  
FT TURN 136 137  
FT STRAND 141 148  
FT STRAND 158 162  
FT TURN 163 164

FT STRAND 165 166  
FT STRAND 175 178  
FT STRAND 183 190  
FT HELIX 193 197  
FT TURN 198 199  
FT STRAND 202 206  
FT STRAND 215 219  
FT STRAND 227 227  
FT STRAND 230 234  
FT STRAND 238 240  
FT TURN 241 242  
FT STRAND 245 256  
FT STRAND 260 266  
FT TURN 267 268  
FT STRAND 269 270  
FT STRAND 274 276  
FT TURN 280 281  
FT TURN 283 284  
FT STRAND 287 296  
FT HELIX 297 301  
FT TURN 302 303  
FT STRAND 306 312  
FT TURN 313 314  
FT TURN 316 317  
FT STRAND 320 324  
SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
  
Query Match 55.1%; Score 1234; DB 1; Length 330;  
Best Local Similarity 92.1%; Pred. No. 6.1e-77;  
Matches 232; Conservative 4; Mismatches 6; Indels 10; Gaps 2;  
  
QY 159 THDNICS-----GNSESQKV-----DKTHTCCPPCAPPELLGGPSVFLFPKPKDTLMIS 208  
DB 78 TQTYICNVNHPSTKVDKVPKSCDKTHTCCPPCAPPELLGGPSVFLFPKPKDTLMIS 137  
QY 209 RTPVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVSVLTVHLQDWL 268  
DB 138 RTPVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVSVLTVHLQDWL 197  
QY 269 NGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYF 328  
DB 198 NGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYF 257  
QY 329 SDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCSCVMHEALHN 388  
DB 258 SDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQQGNVFCSCVMHEALHN 317  
QY 389 HVTQKSLSLSPG 400  
DB 318 HVTQKSLSLSPG 329  
  
RESULT 2  
GC2\_HUMAN  
ID GC2\_HUMAN STANDARD; PRT; 326 AA.  
AC P01859;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2 CHAIN C REGION.  
GN IGHG2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=82197621; PubMed=6804948;  
RA Ellison J.W., Hood L.E.;  
RT "Linkage and sequence homology of two human immunoglobulin gamma  
RT heavy chain constant region genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).  
RN [2]



SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
RX MEDLINE=81007873; PubMed=6774012;  
RA Wang A.-C., Tung E., Fudenberg H.H.;  
RT "The primary structure of a human IgG2 heavy chain: genetic,  
evolutionary, and functional implications.";  
RL J. Immunol. 125:1048-1054(1980).  
[3]  
RN SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).  
RX MEDLINE=80001357; PubMed=113060;  
RA Connell G.E., Parr D.M., Hofmann T.;  
RT "The amino acid sequences of the three heavy chain constant region  
domains of a human IgG2 myeloma protein.";  
RL Can. J. Biochem. 57:758-767(1979).  
[4]  
RN SEQUENCE OF 238-275 (ZIE).  
RX MEDLINE=80114419; PubMed=118920;  
RA Hofmann T., Parr D.M.;  
RT "A note of the amino acid sequence of residues 381-391 of human  
immunoglobulins gamma chains.";  
RL Mol. Immunol. 16:923-925(1979).  
[5]  
RN REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).  
RA Hofmann T., Parr D.M.;  
RL Submitted (MAR-1980) to the PIR data bank.  
[6]  
RN SEQUENCE OF 1-121 (DOT).  
RX MEDLINE=9525298; PubMed=7737190;  
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;  
RT "Characterization of the two unique human anti-flavin monoclonal  
immunoglobulins.";  
RL Eur. J. Biochem. 228:886-893(1995).  
[7]  
RN DISULFIDE BONDS.  
RX MEDLINE=72033500; PubMed=4940472;  
RA Milstein C., Frangione B.;  
RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";  
RL Biochem. J. 121:217-225(1971).  
[8]  
RN DISULFIDE BONDS.  
RX MEDLINE=69064124; PubMed=5782707;  
RA Frangione B., Milstein C., Pink J.R.L.;  
RT "Structural studies of immunoglobulin G.";  
RL Nature 221:145-148(1969).  
[9]  
RN This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
[10]  
RN EMBL; V00554; CAA23814.1; -;  
DR EMBL; V00554; CAA23815.1; -;  
DR EMBL; V00554; CAA23816.1; -;  
DR EMBL; V00554; CAA23817.1; -;  
DR PIR; A02148; G2HU.  
DR MTM; 147110; -;  
DR INTERPRO; IPR000495; -;  
DR INTERPRO; IPR003006; -;  
DR PFAM; PF00047; Ig; 3.  
DR PROSITE; PS00290; IG\_MHC; 2.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1  
FT DOMAIN 1 98 CH1.  
FT DOMAIN 99 110 HINGE.  
FT DOMAIN 111 219 CH2.  
FT DOMAIN 220 326 CH3.  
FT INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 14 14  
FT DISULFID 27 83  
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 140 200  
FT DISULFID 246 304  
FT VARIANT 60 60  
FT SITE 156 156 /FTID=VAR\_003889.  
FT MOD\_RES 326 326 AT OR NEAR THE COMPLEMENT-BINDING SITE.  
SQ SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;  
Query Match 51.3%; Score 1148.5; DB 1; Length 326;  
Best Local Similarity 90.3%; Pred. No. 3.7e-71;  
Matches 214; Conservative 8; Mismatches 8; Indels 7; Gaps 2;  
QY 170 STOKVDKT-----HTCPGPCPAPELLGGPSVFLFPPPKDFTLMISRPETVCVVVDVSHE 223  
DB 90 SNTKVDKTVVERKCCVECPAPP-VAGPSVFLFPPPKDFTLMISRPETVCVVVDVSHE 148  
QY 224 DPEVKFNWYDGVGEVHNAKTKPREEQYNSTYRVVSVLTVLHODWLNKGKEYCKVSNKALP 283  
DB 149 DPEVQFNWYDGVGEVHNAKTKPREEQFNSTFRVSVLTVVHQDWLNKGKEYCKVSNKGLP 208  
QY 284 APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGOPEN 343  
DB 209 APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGOPEN 268  
QY 344 NYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNHYTQKSLSLSPG 400  
DB 269 NYKTPPMLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEALHNHYTQKSLSLSPG 325  
RESULT 3  
GC4\_HUMAN  
ID GC4\_HUMAN STANDARD; PRT; 327 AA.  
AC P01861;  
DT 21-JUL-1986 (Rel. 01; Created)  
DT 21-JUL-1986 (Rel. 01; Last sequence update)  
DT 15-JUL-1999 (Rel. 38; Last annotation update)  
DE IG GAMMA-4 CHAIN C REGION.  
GN IGHG4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83157104; PubMed=6299662;  
RA Ellison J.W., Buxbaum J.N., Hood L.E.;  
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";  
RL DNA 1:11-18(1981).  
RN [2]  
RP SEQUENCE OF 1-30 AND 81-326.  
RX MEDLINE=70207560; PubMed=4192699;  
RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;  
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the  
constant region of a gamma 4 chain";  
RL Biochem. J. 117:33-47(1970).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
[11]  
RN EMBL; K01316; AAB59394.1; ALT\_INIT.  
DR EMBL; K01316; AAB59394.1; ALT\_INIT.  
DR PIR; A02150; G4HU.  
DR MTM; 147130; -;  
DR INTERPRO; IPR000495; -;  
DR INTERPRO; IPR003006; -;  
DR PFAM; PF00047; Ig; 3.  
DR PROSITE; PS00290; IG\_MHC; 2.  
KW Immunoglobulin domain; Immunoglobulin C region.



SQ SEQUENCE 290 AA; 32331 MW; E69C95705B2F46 CRC64;  
 Query Match 50.6%; Score 1133.5; DB 1; Length 290;  
 Best Local Similarity 70.88; Pred. No. 3.4e-70;  
 Matches 218; Conservative 19; Mismatches 22; Indels 49; Gaps 5;  
 QY 97 CLKHRS-PPGFGVVQAGTPERNTVCKRCPCDGGFFNETSSKAP-CRKHTNCSVFGLL 152  
 Db 27 CPEPKSCDTPPP-----CPRCPEP-KSODTPPCPCPEPKSC----- 63  
 QY 153 TOKGNATHDNCSNSEQSTOKVTHCTPCPCAPBELLGGPSVFLPPKPKDTLMISRTPE 212  
 Db 64 -----DTPPCPCPCAPBELLGGPSVFLPPKPKDTLMISRTPE 101  
 QY 213 VTCVVVDVSHEDPVKFNWYVDGVEVHNKTKPREQYNSYRVSVLTVLHQDLNKGKE 272  
 Db 102 VTCVVVDVSHEDPEVQFKNWYVDGVVHNKTKPREQYNSYRVSVLTVLHQDLNKGKE 161  
 QY 273 YKCVSNKALPAPIEKTISKAKGQPREPQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIA 332  
 Db 162 YKCVSNKALPAPIEKTISKAKGQPREPQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIA 221  
 QY 333 VEWESGQPNNTYKTPPVLDSGFSFLYKSLTVDKSRWQGNVFCSCVMHEALHNHYTQ 392  
 Db 222 VEWESGQPNNTYKTPPVLDSGFSFLYKSLTVDKSRWQGNVFCSCVMHEALHNHYTQ 281  
 QY 393 KSLSLSPG 400  
 Db 282 KSLSLSPG 289  
 RESULT 5  
 GC\_RABIT STANDARD; PRT; 323 AA.  
 AC POL1870;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA CHAIN C REGION.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84030930; PubMed=6313520;  
 RA Bernstein K.E., Alexander C.B., Mage R.G.;  
 RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant  
 F-1 haplotype."  
 RL Immunogenetics 18:387-397(1983).  
 RN [2]  
 RP SEQUENCE OF 1-128.  
 RX MEDLINE=76135469; PubMed=1243651;  
 RA Pratt D.M., Mole L.E.;  
 RT "Sequence studies on the constant region of the Fd sections of rabbit  
 immunoglobulin G of different allotype."  
 RL Biochem. J. 151:337-349(1975).  
 RN [3]  
 RP SEQUENCE OF 88-266 FROM N.A.  
 RX MEDLINE=83299917; PubMed=6193512;  
 RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;  
 RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma  
 heavy chain and identification of two genomic C gamma genes."  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).  
 RN [4]  
 RP SEQUENCE OF 132-161.  
 RX MEDLINE=70110015; PubMed=5461106;  
 RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;  
 RT "Sequence studies of the Fd section of the heavy chain of rabbit  
 immunoglobulin G."  
 RL Biochem. J. 116:249-259(1970).  
 RN [5]  
 RP SEQUENCE OF 129-131 AND 155-322.

RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;  
 RL (In) Killander J. (eds.);  
 RL Gamma globulins, Nobel symp. 3, pp.109-127, Almquist and Wiksell,  
 Stockholm (1967).  
 CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,  
 CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15  
 CC MARKERS AND REF.5 THE E15 MARKER.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M16426; AAA31289.1; -;  
 DR PIR; A02161; GHRB.  
 DR INTERPRO; IPR000495; -;  
 DR INTERPRO; IPR003006; -;  
 DR PFAM; PF00047; Ig; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1  
 FT VARIANT 104 104 T -> M (IN D11 MARKER).  
 FT VARIANT 185 185 T -> A (IN E15 MARKER).  
 FT CONFLICT 48 48 N -> E (IN REF. 2).  
 FT CONFLICT 71 71 V -> VPV (IN REF. 2).  
 FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).  
 FT CONFLICT 173 173 N -> D (IN REF. 5).  
 FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).  
 FT CONFLICT 201 201 N -> D (IN REF. 5).  
 FT CONFLICT 218 218 Q -> E (IN REF. 5).  
 FT CONFLICT 233 233 E -> Q (IN REF. 5).  
 FT CONFLICT 246 246 N -> D (IN REF. 5).  
 FT CONFLICT 256 256 E -> G (IN REF. 5).  
 FT CONFLICT 260 260 N -> D (IN REF. 5).  
 FT CONFLICT 266 266 Q -> W (IN REF. 5).  
 FT CONFLICT 280 280 Y -> W (IN REF. 5).  
 FT CONFLICT 284 284 N -> S (IN REF. 5).  
 SQ SEQUENCE 323 AA; 35404 MW; 69E8A118D579A8B CRC64;  
 Query Match 41.2%; Score 923; DB 1; Length 323;  
 Best Local Similarity 61.58; Pred. No. 6.7e-56;  
 Matches 177; Conservative 33; Mismatches 46; Indels 32; Gaps 5;  
 QY 123 RCPDGFSS-----NETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCSNSEQSTOKVDTK 177  
 Db 57 RQSSGLYSLSSVSVTSSSQP-----VTCNV-----AHPATNFKVDKT 94  
 QY 178 ---HTC--PPCAPBELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFWNY 232  
 Db 95 VAPSTCSKPTCPPELGGPSVFIFFPKPKDTLMISRTPEVTCVVVDVSDDEPEYQFTWY 154  
 QY 233 VDGVEVHNKTKPREQYNSYRVSVLTVLHQDLNKGKCKVSNKALPAPIEKTISK 292  
 Db 155 INNEQVTRARPLREQQNFNSTIRVVSTLPIHQDLNKGKCKVSNKALPAPIEKTISK 214  
 QY 293 AKGQPREPQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESGQPNNTYKTPPV 352  
 Db 215 ARGQPLEPKVYTMGPPREELSSRSVSLTCMNGFYPSDISVEWESGQPNNTYKTPPV 274  
 QY 353 DSDGSFFLYKSLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSPG 400  
 Db 275 DSDGSFFLYKSLVPTSEWQRGDVFCTCSVMHEALHNHYTQKSLSPG 322  
 RESULT 6  
 GC2\_CAVPO  
 ID GC2\_CAVPO STANDARD; PRT; 329 AA.  
 AC P01862;  
 DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-2 CHAIN C REGION.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 RN [1]  
 RP SEQUENCE OF 1-3.  
 RA Trischmann T.M.;  
 RL Submitted (APR-1975) to the PIR data bank.  
 RN [2]  
 RP SEQUENCE OF 4-68.  
 RX MEDLINE=71058471; PubMed=5538606;  
 RA "Structure of heavy chain from strain 13 guinea pig  
 RT "Structure of heavy chain from strain 13 guinea pig  
 RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal  
 RT and hinge region cyanogen bromide fragments.";  
 RL Biochemistry 10:9-17(1971).  
 RN [4]  
 RP SEQUENCE OF 134-226.  
 RX MEDLINE=75036072; PubMed=4429665;  
 RA Tracey D.E., Cebra J.J.;  
 RT "Primary structure of the CH2 homology region from guinea pig IgG2  
 RT antibodies.";  
 RL Biochemistry 13:4796-4803(1974).  
 RN [5]  
 RP SEQUENCE OF 227-311.  
 RX MEDLINE=75036073; PubMed=4609467;  
 RA Trischmann T.M., Cebra J.J.;  
 RT "Primary structure of the CH3 homology region from guinea pig IgG2  
 RT antibodies.";  
 RL Biochemistry 13:4804-4811(1974).  
 RN [6]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=71058474; PubMed=4922544;  
 RA Oliveira B., Lamm M.E.;  
 RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";  
 RL Biochemistry 10:26-31(1971).  
 CC 1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN  
 CC 13 INBRED GUINEA PIGS.  
 DR PIR; A02151; G2GP.  
 DR INTERPRO; IPR000495; -.  
 DR INTERPRO; IPR003006; -.  
 DR PFAM; PF00047; Ig; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 16 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 16 16  
 FT DISULFID 28 79  
 FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 142 202  
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).  
 FT DISULFID 248 308  
 SQ SEQUENCE 329 AA; 36074 MW; 50231B7164DIFBA9 CRC64;  
 Query Match 40.4%; Score 904.5; DB 1; Length 329;  
 Best Local Similarity 60.3%; Pred. No. 1.2e-54;  
 Matches 173; Conservative 30; Mismatches 53; Indels 31; Gaps 3;  
 QY 127 GFSSNETSKAPCRKHTNCSVFGLLLTQGNATHDNCISGNSSTQKVDKT----- 177  
 DB 61 GLYSLTSMVTVPSSQKATCNV-----AHPASSTKVDKTVPIRT2P 102

QY 178 --HTCPCPAPELLGGPSVFLPPPKPDKTLMISRTPEVTCVVVDVSHEDPEVKFNWYDG 235  
 DB 103 BCTCPKCPPPENLGGPSVFIFFPKPKDTLMISLTSLPRVTCVVVDVSDDEPEVFQTFWVDN 162  
 QY 236 VEVHNAKTPREEQYNSYRVSVTLVHQLDNLNGKEYCKYKSNKALPAPIEKTISKAKG 295  
 DB 163 KPGVNAETKPRVEQNTTFRVESLPIQHODWLRGKFKCKYKSNKALPAPIEKTISKAKG 222  
 QY 296 OPREPQVYTLPPSRDELTKNOVSLTCLVKGFGYPSDIATVWESNGQP--ENNKTTPPVLD 353  
 DB 223 APRMPDVYTLPPSRDELSKSVSVTCLINFFPADIHVWASNRVPSVSEKYNTPPIED 282  
 QY 354 SDGSEFLYSKLTVDKSRWQGVNFSCSVNHEALHNHYTQKSLSLSPG 400  
 DB 283 ADGSYFLYSKLTVDKSAWDQGTIVTCSVMHEALHNHYTQKSLSRSPG 329  
 RESULT 7  
 GC3\_MOUSE STANDARD; PRT; 329 AA.  
 AC P22436;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-3 CHAIN C REGION, SECRETED FORM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85027161; PubMed=6092053;  
 RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,  
 RA Tucker P.W., Blattner F.R.;  
 RT "Structural analysis of the murine IgG3 constant region gene.";  
 RL EMBO J. 3:2041-2046(1984).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; J00451; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; B02156; G3MSC.  
 DR INTERPRO; IPR000495; -.  
 DR INTERPRO; IPR003006; -.  
 DR PFAM; PF00047; Ig; 3.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Immunoglobulin C region; Glycoprotein; Transmembrane;  
 KW Alternative splicing.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 223 CH2.  
 FT DOMAIN 224 327 CH3.  
 SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;  
 Query Match 37.8%; Score 847.5; DB 1; Length 329;  
 Best Local Similarity 54.0%; Pred. No. 8.9e-51;  
 Matches 162; Conservative 47; Mismatches 72; Indels 19; Gaps 5;

QY 107 FGWVQAGTPERTVCKPCDGFSSNETSKAPCRKHTNCSVFGLLLTQGNATHDNCISG 166  
 DB 42 YGALSSGV---RTVSSVLQSGFYSLSLVTPVSTWPSQTIVT-----CNVAHF---AS 88  
 QY 167 NSESTQKVD-----KTHTCP--PCPAPELLGGPSVFLPPPKPDKTLMISRTPEVTCVVVDV 220  
 DB 89 KTELKRIEPRIPKPSPTPGSSCPGNILGGPSVFIFPPPKPKDALMISLTPTKVTCCVVVDV 148

```

QY 221 SHEDPEVKFNWYDGVGVHNAKTPREEQNSTYRVVSVLTVLHQLDNLNGKEYCKKYSNK 280
D 149 SEDDPDVHVSFVNDKENVHTAWTPREAOQNSTFRVSALPIQHDWNRGKEFKCKVNNK 208
QY 281 ALPAPIEKTISKAGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWESNGQ 340
D 209 ALPAPIERTISKGRATQPVYTIIPPRQMSKKVSLTCLVTFNFSEAISVEWENGE 268
QY 341 PENNYKTPPVLDSDGFFLYSLTKVDKSRWQOGNVFSCSVHHEALHNHYTKLSLSLSPG 400
D 269 LEQDYKNTPVLDSDGFFLYSLTKVDKSRWQOGNVFSCSVHHEALHNHYTKLSLSLSPG 328

RESULT 8
GCB_RAT
ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2B CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89232738; PubMed-3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR INTERPRO: PS0018;
DR INTERPRO: IPR000495;
DR PFAM: PF00047; ig; 3;
DR PROSITE: PS00290; IG_MHC; 1;
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 37.6%; Score 843; DB 1; Length 333;
Best Local Similarity 56.3%; Pred. No. 1.8e-50;
Matches 161; Conservative 38; Mismatches 53; Indels 34; Gaps 4;

QY 130 SNETSSKAPCRKHTNCVSFGLLLTQKGNATHDNCNSSESTQKVDK-----HTC 180
D 66 SSVTSWPSQTVT-CNV-----AHPASSTKVDKVRNGGIGHC 106

QY 181 P-----PCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVYVDVSHEDPEVKFNWYD 234
D 107 PTCPTCHKCPPELLGGPSVFLPPKPKDTLLISQNAKVCVYVDVSEEDPVQVSVFN 166

QY 235 GVEVHNAKTPREEQNSTYRVVSVLTVLHQLDNLNGKEYCKKYSNKALPAPIEKTISKAK 294
D 167 NVEVHTAQTPREEQNSTYRVVSVLTVLHQLDNLNGKEYCKKYSNKALPSPKTIKPK 226

QY 295 GQREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWESNGQENYKTPPVLD 354
D 227 GLVRKPKQVYMGPPTEQLTQVSLTCLTSGFLPNDIGVEWTSNGHTEKNYKNTPEVMD 286

QY 355 DGSFELYSLTKVDKSRWQOGNVFSCSVHHEALHNHYTKLSLSLSPG 400
D 287 DGSEFFMYSKLNVERSRDSDRAPFCVSVHHEALHNHYTKLSLSLSPG 332

```

```

RESULT 9
GCB_MOUSE
ID GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-85027161; PubMed-6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
[2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE-84041483; PubMed-6314258;
RA Komaromy M., Clayton L., Rogers J., Roberts S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; J00451; AAB59655.1;
DR EMBL; V01526; CAA24767.1; ALT_SEQ.
DR PIR; A02155; G3MSM.
DR INTERPRO: IPR000495;
DR INTERPRO: IPR003006;
DR PFAM: PF00047; ig; 3;
DR PROSITE: PS00290; IG_MHC; 1;
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 98 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match 37.6%; Score 841.5; DB 1; Length 398;
Best Local Similarity 53.8%; Pred. No. 2.8e-50;
Matches 161; Conservative 47; Mismatches 72; Indels 19; Gaps 5;

QY 107 FGVOAGTPEKRTVCKPCDGFSSNETSSKAPCRKHTNCVSFGLLLTQKGNATHDNC 166
D 42 YGALSSV-----RTVSSVLSQSGFVSLSLTVTPSSWPSQTVI-----CNVAP 88

QY 167 NSESTQKVD-----KTHTCP--PCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVYVDV 220
D 89 KTELKRIEPRIPKSTPPGSSCPPGNGILGGPSVFIIPKPKDALMSLTPKTCVYVDV 148

QY 221 SHEDPEVKFNWYDGVGVHNAKTPREEQNSTYRVVSVLTVLHQLDNLNGKEYCKKYSNK 280
D 149 SEDDPDVHVSFVNDKENVHTAWTPREAOQNSTFRVSALPIQHDWNRGKEFKCKVNNK 208

```



```
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 102 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45E49B9DA CRC64;

Query Match 36.6%; Score 820; DB 1; Length 326;
Best Local Similarity 60.5%; Pred. No. 6.4e-49;
Matches 146; Conservative 42; Mismatches 39; Indels 14; Gaps 3;

QY 170 STOKVDKTH-----HTCRPPAPPELLGG---PSVLFPPPKDITLMISRTPEVTCVVVD 219
Db 89 SSTKVDKKIVPRNCGGDKPC----ICTGEVSVFIFPPKPDVLTITLTPKVCVVVD 144
QY 220 VSHEDPEVKFNWYDGVGVHNAKTPREEQNSTYRVSVTLVHQLDNLNGKEYCKVSN 279
Db 145 ISQDDPEVHFSWFVDVVEVHTAOTRPEEQFNSTFRSVSELPILHQLDNLNGTRFCRCKVTS 204
QY 280 KALPAIEKTIKAKGPREPQVYITLPPSRDELTKNQVSLTCLVKGPYPSDIAVEWESNG 339
Db 205 AAPFSPIEKTIKSGRTOVPHYVTMSPTKEEMTQNEVSTCMVKGFYPPDIYVWQMNG 264
QY 340 OPENNYKTPPVLDSDGSFELYSLKLTVDKSRWQGVNFCVSMHEALHNHYTKSLSP 399
Db 265 QPENYKNTPTMTDGSVFLYSLKLVKKEKQOQNTFTCSVLHGLNHNHTEKSLSHSP 324
QY 400 G 400
Db 325 G 325

RESULT 12
GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
[2]
SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowan A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
[3]
SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
[4]
SEQUENCE OF 1-44 FROM N.A.
```

```
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V00793; CAA241172.1; -
DR EMBL; V00793; CAA241173.1; -
DR EMBL; V00793; CAA241174.1; -
DR PIR; B02159; GIMSM.
DR MGI; 96446; IGH-4.
DR INTERPRO; IPR000495; -
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; Igh; 3.
DR PROSITE; PS00290; IGH_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 HINGE.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC... ).
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393 POTENTIAL.
FT SEQUENCE 393 AA; 43386 MW; 4CC88343B7AICE27 CRC64;
SQ

Query Match 36.6%; Score 820; DB 1; Length 393;
Best Local Similarity 60.5%; Pred. No. 7.9e-49;
Matches 144; Conservative 45; Mismatches 39; Indels 10; Gaps 3;

QY 170 STOKVDKTH-----CPP--CPAPPELLGGPSVFLFPPPKDITLMISRTPEVTCVVVDVSH 222
Db 89 SSTKVDKKIVPRDCGCKPCICTVPEV---SSVFIFPPKPDVLTITLTPKVCVVVDISK 145
QY 223 EDEVEKFNWYDGVGVHNAKTPREEQNSTYRVSVTLVHQLDNLNGKEYCKVSNKAL 282
Db 146 DDEVEQVSWFVDDVVEVHTAOTRPEEQFNSTFRSVSELPINHQLDNLNGKEYCKRVNSAAF 205
QY 283 PAPEKTIKAKGPREPQVYITLPPSRDELTKNQVSLTCLVKGPYPSDIAVEWESNSQPE 342
Db 206 PAPEKTIKTKGRKAPQVITPPPREQMAKQKQVITMTITDFFEDITVWQWNGQPA 265
QY 343 NNYKTPPVLDSDGSFFLYSLKLTVDKSRWQGVNFCVSMHEALHNHYTKSLSPG 400
Db 266 ENYKNTQPIINTNGSYFYSLKLVKQVNSWEAGNTFTCSVLHGLNHNHTEKSLSHSPG 323

RESULT 13
GCAA_MOUSE STANDARD; PRT; 330 AA.
ID GCAA_MOUSE
```



AC P01863;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE IG GAMMA-2A CHAIN C REGION, A ALLELE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=81076554; PubMed=6777755;  
RA Sikorav J.-L., Auffray C., Rougeon F.;  
RT "Structure of the constant and 3' untranslated regions of the murine  
RT Balb/c gamma 2a heavy chain messenger RNA.";  
RL Nucleic Acids Res. 8:3143-3155(1980).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=81198976; PubMed=6267279;  
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;  
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene  
RT and evolution of heavy chain genes: further evidence for intervening  
RT sequence-mediated domain transfer.";  
RL Nucleic Acids Res. 9:1365-1381(1981).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=81223894; PubMed=6787604;  
RA Ollo R., Auffray C., Morchamps C., Rougeon F.;  
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes  
RT suggests that exons can be exchanged between genes in a multigenic  
RT family.";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).  
RN [4]  
RP MYELOMA PROTEIN MOPC 173.  
RX MEDLINE=74175517; PubMed=4831970;  
RA Bourgois A., Fougereau M., Rocca-Serra J.;  
RT "Determination of the primary structure of a mouse IgG2a  
RT immunoglobulin: amino-acid sequence of the Fc fragment. Implications  
RT for the evolution of immunoglobulin structure and function.";  
RL Eur. J. Biochem. 43:423-435(1974).  
RN [5]  
RP DISULFIDE BONDS.  
RX MEDLINE=73056887; PubMed=4565406;  
RA de Preval C., Fougereau M.;  
RT "Determination of the primary structure of a mouse gamma G2a  
RT immunoglobulin. Identification of the disulfide bridges.";  
RL Eur. J. Biochem. 30:452-462(1972).  
RN  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; V00798; CAA24178.1; -  
CC PIR; A02152; G2MSA.  
CC INTERPRO; IPR000495; -  
CC INTERPRO; IPR003006; -  
CC PFAM; PF000047; Ig; 3.  
CC PROSITE; PS00290; IG\_MHC; 1.  
CC Immunoglobulin domain; Immunoglobulin C region.  
CC NON\_TER 1 1  
CC DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
CC DISULFID 27 82  
CC DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
CC DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
CC DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
CC DISULFID 144 204  
CC DISULFID 250 308  
CC DISULFID 330 330 REMOVED POST-TRANSLATIONALLY.  
CC SEQUENCE 330 AA; 36389 MW; B84361C545A6864 CRC64;

Query Match 36.3%; Score 812.5; DB 1; Length 330;  
Best Local Similarity 56.5%; Pred. No. 2.1e-48;  
Matches 157; Conservative 34; Mismatches 58; Indels 29; Gaps 4;  
Qy 133 TSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSSTQKVDK-----THTCPP-- 182  
Db 71 TSSWPSQSI-CNV-----AHPASSTKVDKKEIPRGPTIKPCPK 111  
Qy 183 CPAPELLGGSPVFLPPPKDTLMSRTPETCVVVDVSHEDPEVKFWYVDGVVHNK 242  
Db 112 CPAPNLLGGSPVFLPPPKDVLMSLSPITCVVVDVSEDDPDVQISWFVNNVEVHTAQ 171  
Qy 243 TKPREQYNSYRYVSVLTVLHQDWLNGKEYKCKVCKSKALPAPIETKTSKAGOPREPQV 302  
Db 172 TQTHREDYNSLRYVVSALPIQHDWMSGKEFKCKVNNKDLPAPIETKSKAGSVRAPOV 231  
Qy 303 YTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYS 362  
Db 232 YVLPPEPEETKKQVTLTCMVTDMPEDYVVENTNGKTELNYKNTPEVLDSDGSFYMS 291  
Qy 363 KLTVDKSRWQOGNVSFSCSWMHEALHNHYTKQSLSPG 400  
Db 292 KLRVEKKNNWERNYSYSCSVWHEGLHNHHTTKFSRTPG 329  
RESULT 14  
GCAM\_MOUSE STANDARD; PRT; 399 AA.  
ID GCAM\_MOUSE AC P01865;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2A CHAIN C REGION, MEMBRANE-BOUND FORM.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=82222190; PubMed=6283537;  
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;  
RT "Nucleotide sequences of gene segments encoding membrane domains of  
RT immunoglobulin gamma chains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).  
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGS CONTAIN TWO MRNA  
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED  
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-  
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED  
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND  
CC SEGMENT OF MU CHAINS.  
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE  
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF  
CC THE A ALLELE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J00471; AAB59661.1; ALT\_INIT.  
CC PIR; A02154; G2MSA.  
CC MGD; MGI:96443; IGH-1.  
CC INTERPRO; IPR000495; -  
CC INTERPRO; IPR003006; -  
CC PFAM; PF000047; Ig; 3.  
CC PROSITE; PS00290; IG\_MHC; 1.  
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
CC Transmembrane; Alternative splicing.  
CC NON\_TER 1 1  
CC DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).



```
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363
FT DOMAIN 364 399
FT CARBOHYD 180 180
SQ SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;

Query Match 36.3%; Score 812.5; DB 1; Length 399;
Best Local Similarity 56.5%; Pred. No. 2.6e-48;
Matches 157; Conservative 34; Mismatches 58; Indels 29; Gaps 4;

QY 133 TSSKAPCRKHTNCSVFGLLLTQKGNATHDNCNSSESTOKVDK-----RHTCSP-- 182
Db 71 TSSTWPSQSIT-CNV-----AHPASSTKVDDKIEPRGPTIKPCPPCK 111
QY 183 CPAPELLGGPSVFLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 242
Db 112 CPAPNLLGGPSVFLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAK 171
QY 243 TKPREQYNSTYRVVSVLTVLHODWLNKGEYCKVSKNPKALPAPIETISKAKGP 302
Db 172 TQTHREDYNSTLRVVSALPIQHQDWSGKGEFKCKVNNKDLPAPIETISKPKGSVRAPQV 231
QY 303 YTLPPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSF 362
Db 232 YVLPPEEEMTKQVLTCTMTVDMPEDYVETWNNKTELYNKTEPVLDSDGSYFMY 291
QY 363 KLVTDKSRWQGNVSCSVMHALHNHYTKLSLSPG 400
Db 292 KLRVEKNWYERNYSYSCSVVHGLHNHYTKSPRTFC 329

RESULT 15
GCC_RAT GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2C CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88166903; PubMed-312722;
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RL region cDNA: extensive homology to mouse gamma 3."
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07189; CAA30169.1; -
DR PIR; S00847; S00847.
DR INTERPRO; IPR000495; -
DR INTERPRO; IPR003006; -
DR PFAM; PF000047; ig. 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
```

```
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B793850773 CRC64;

Query Match 35.9%; Score 804.5; DB 1; Length 329;
Best Local Similarity 63.7%; Pred. No. 7.2e-48;
Matches 142; Conservative 41; Mismatches 37; Indels 3; Gaps 1;

QY 181 PP---CPAPELLGGPSVFLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 237
Db 106 PPTDICSCDDNLGRPSVFIFPPKPKDILMITLPKVTCTVVVDVSEEPDQVQSFVDMVR 165
QY 238 VHNAKTKPREOYNSTYRVVSVLTVLHODWLNKGEYCKVSKNPKALPAPIETISKAKGP 297
Db 166 VFTAQTPHEQLNGTFRVWSTLHIQHDMMSGREFCKVNNKDLPSPIETISKPRGKA 225
QY 298 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGS 357
Db 226 RTPQVYTIPTPREQMSKNKVSITCMVTSFYFASISVEWERNGELEQDYKNTLPVLDSDS 285
QY 358 FFLYSKLTVDKSRWQGNVSCSVMHALHNHYTKLSLSPG 400
Db 286 YFLYSKLSVDTDSWNRGDIYTCVVHEALHNHHTQKNLSRSPG 328
```

Search completed: March 1, 2001, 09:17:45  
Job time: 263 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:45 ; Search time 79.26 Seconds  
(without alignments)  
342.673 Million cell updates/sec

Title: US-09-389-782A-5  
Perfect score: 2240  
Sequence: 1 ETFFPKYLHYDEETSHQLLC.....VMHEALHNHYTKSLSPG 400

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results, predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1234	55.1	330	1	GHU	Ig gamma-1 chain C
2	1230	54.9	374	2	S72664	Ig heavy chain V r
3	1229	54.9	255	4	S31866	Ig gamma-1 chain C
4	1225	54.7	374	2	S69339	Ig heavy chain V r
5	1183.5	52.8	234	2	PT0207	Ig gamma chain C r
6	1154.5	51.5	377	2	A23511	Ig gamma-3 chain C
7	1152.5	51.5	377	2	A60764	Ig gamma-3 chain C
8	1148.5	51.3	326	1	G2HU	Ig gamma-2 chain C
9	1136	50.7	327	1	G4HU	Ig gamma-2 chain C
10	1133.5	50.6	289	1	G3H0W1	Ig gamma-4 chain C
11	923	41.2	323	1	GHRB	Ig gamma-3 heavy c
12	917.5	41.0	328	2	I47160	Ig gamma chain C r
13	917.5	41.0	328	2	I47159	Ig gamma 2b chain
14	915.5	40.9	277	2	I47162	Ig gamma 2a chain
15	904.5	40.4	329	1	G2GP	Ig gamma 4 chain c
16	895.5	40.0	328	2	I47158	Ig gamma-2 chain C
17	889.5	39.7	328	2	I47161	Ig gamma 1 chain C
18	861.5	38.5	470	2	S22080	Ig gamma 3 chain c
19	851.5	38.0	308	2	G30554	Ig heavy chain pre
20	851.5	38.0	472	2	S31459	Ig heavy chain C r
21	847.5	37.8	329	1	G3MSC	Ig gamma-1 chain -
22	843	37.6	333	2	PS0018	Ig gamma-3 chain C
23	841.5	37.6	398	1	G3MSM	Ig gamma-2b chain
24	830	37.1	444	2	PC4436	Ig gamma-3 chain C
25	820	36.6	324	1	G1MS	monoclonal antibody
26	820	36.6	326	2	PS0017	Ig gamma-1 chain C
27	820	36.6	393	1	G1MSM	Ig gamma-1 chain C
28	812.5	36.3	330	1	G2MSA	Ig gamma-2a chain
29	812.5	36.3	399	1	G2MSAM	Ig gamma-2a chain

## ALIGNMENTS

RESULT 1

GHU

Ig gamma-1 chain C region - human

C;Species: Homo sapiens (man)

C;Date: 31-Jan-1981 #sequence.revision 18-Aug-1982 #text.change 16-Jul-1999

C;Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R;Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A;Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A;Reference number: A93433; MUID:82274238

A;Accession: A93433

A;Molecule type: DNA

A;Residues: 1-330 <ELL>

A;Cross-references: EMBL:Z17370

A;Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) marker

A;Note: Lys-330 is removed after translation

R;Harris, L.J.

submitted to the EMBL Data Library, October 1992

A;Reference number: S33904

A;Accession: S36861

A;Molecule type: DNA

A;Residues: 2-330 <HAR>

A;Cross-references: EMBL:Z17370

R;Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A;Title: Structure of human immunoglobulin gamma genes: implications for evolution of

A;Reference number: S33887; MUID:83001943

A;Accession: S33887

A;Molecule type: DNA

A;Residues: 88-113;235-330 <TAK>

A;Cross-references: EMBL:Z17370

R;Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman

Biochemistry 9, 3161-3170, 1970

A;Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq

A;Reference number: A90563; MUID:71064024

A;Contents: myeloma protein Eu

A;Accession: B90563

A;Molecule type: protein

A;Residues: 1-96,'R',98-135 <CUN>

A;Note: this sequence has the G1m(3) marker, 97-Arg

R;Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A;Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se

A;Reference number: A90564; MUID:71064025

A;Contents: Eu

A;Accession: A90564

A;Molecule type: protein

A;Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2

A;Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met

R;Ponstingl, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976

A;Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein NI



Db 145 SKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP 204  
QY 351 VLSDSGSFFLYSKLTVDKSRWQOQGVFSCSVMHAEALHNHYTQKSLSLSPG 400  
Db 205 VLSDSGSFFLYSKLTVDKSRWQOQGVFSCSVMHAEALHNHYTQKSLSLSPG 254

## RESULT 4

S69339  
Ig heavy chain V region precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 17-Mar-2000  
C:Accession: S69339  
R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogné, M.  
Eur. J. Biochem. 229, 54-60, 1995  
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.  
A:Reference number: S69339; MUID:95262687  
A:Accession: S69339  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <RHA>  
A:Cross-references: EMBL:X81695  
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 54.7%; Score: 1225; DB 2; Length 374;  
Best Local Similarity 96.6%; Pred. No. 2.7e-68;  
Matches 225; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 168 SESTQKVDKTHTCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEV 227  
Db 141 SSEPSCDKTHTCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEV 200

QY 228 KFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE 287  
Db 201 KFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE 260

QY 288 KTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 347  
Db 261 KTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 320

QY 348 TTPVLDSDGSFFLYSKLTVDKSRWQOQGVFSCSVMHAEALHNHYTQKSLSLSPG 400  
Db 321 TTPVLDSDGSFFLYSKLTVDKSRWQOQGVFSCSVMHAEALHNHYTQKSLSLSPG 373

## RESULT 5

PT0207  
Ig gamma chain C region - chimpanzee  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999  
C:Accession: PT0207  
R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.  
Mol. Immunol. 28, 319-322, 1991  
A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.  
A:Reference number: PT0207; MUID:91287716  
A:Accession: PT0207  
A:Molecule type: mRNA  
A:Residues: 1-234 <EHR>

C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 52.8%; Score: 1183.5; DB 2; Length 234;  
Best Local Similarity 94.4%; Pred. No. 5.7e-66;  
Matches 221; Conservative 1; Mismatches 3; Indels 9; Gaps 1;

QY 170 STQKVDK-----THTCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDV 220  
Db 1 SNTKVDKRVKPKSCDTTHTCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDV 60

QY 221 SHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK 280

Db 61 SHEDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK 120  
QY 281 ALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ 340  
Db 121 ALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ 180  
QY 341 PENNYKTPPVLDSDGSFFLYSKLTVDKSRWQOQGVFSCSVMHAEALHNHYTQKS 394  
Db 181 PENNYKTPPVLDSDGSFFLYSKLTVDKSRWQOQGVFSCSVMHAEALHNHYTQKS 234

## RESULT 6

A23511  
Ig gamma-3 chain C region (allotype G3m(b)) - human  
C:Species: Homo sapiens (man)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999  
C:Accession: A23511  
R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.  
Nucleic Acids Res. 14, 1779-1789, 1986  
A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene.  
A:Reference number: A23511; MUID:86148507  
A:Accession: A23511  
A:Molecule type: DNA  
A:Residues: 1-377 <HUC>  
A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056  
C:Genetics:  
A:Gene: GDB:IGHG3  
A:Cross-references: GDB:119339; OMIM:147120  
A:Map position: 14q32.33-14q32.33  
A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 51.5%; Score: 1154.5; DB 2; Length 377;  
Best Local Similarity 67.1%; Pred. No. 5.7e-64;  
Matches 228; Conservative 19; Mismatches 32; Indels 61; Gaps 7;

QY 76 CNRPHNRVCKEGRYLEIEF-----CLKHRSC---PPGFGVVGAGTPERNTVC 121  
Db 83 CNVNH-KPSNTKVDKRVELKTLGLDTHTCPCRCPEKSCDTPPP-----C 126

QY 122 KRCDGDFESNETSSKAP-CRKHTNCSVFGLLLTQGNATHDNCISGSESTQKVDKTHFC 180  
Db 127 PRCEP-KSCDTPPCRCPEKSC-----DTPPC 156

QY 181 PPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVEN 240  
Db 157 PRCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVQFKWYDGVGVEN 216

QY 241 AKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREP 300  
Db 217 AKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREP 276

QY 301 QVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL 360  
Db 277 QVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFL 336

QY 361 YSKLTVDKSRWQOQGVFSCSVMHAEALHNHYTQKSLSLSPG 400  
Db 337 YSKLTVDKSRWQOQGVFSCSVMHAEALHNHYTQKSLSLSPG 376

## RESULT 7

A60764  
Ig gamma-3 chain C region, form LAT - human  
C:Species: Homo sapiens (man)  
C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 16-Jul-1999  
C:Accession: A60764  
R:Huck, S.; Lefranc, G.; Lefranc, M.P.  
Immunogenetics 30, 250-257, 1989

R; Hofmann, T.; Parr, D.M.  
Mol. Immunol. 16, 923-925, 1979  
A; Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin G  
A; Reference number: A93132; MUID: 80114419  
A; Contents: Zie  
A; Accession: A93132  
A; Molecule type: protein  
A; Residues: 238-275 <HOF>  
R; Hofmann, T.; Parr, D.M.  
submitted to the Atlas, March 1980  
A; Reference number: A94591  
A; Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268  
A; Note: the revised sequence differs from that shown in having 60-Ala and in the amid  
ned  
R; Milstein, C.; Frangione, B.  
Biochem. J. 121, 217-225, 1971  
A; Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.  
A; Reference number: A90253; MUID: 72033500  
A; Contents: annotation; myeloma protein Sa, disulfide bonds  
R; Frangione, B.; Milstein, C.; Pink, J.R.L.  
Nature 221, 145-148, 1969  
A; Title: Structural studies of immunoglobulin G.  
A; Reference number: A93157; MUID: 69064124  
A; Contents: annotation; Sa, disulfide bonds  
C; Genetics:  
A; Gene: GDB:IGHG2  
A; Cross-references: GDB:119338; OMIM:147110  
A; Map position: 14q32.33-14q32.33  
C; Complex: An immunoglobulin heterotetramer subunit consists of two identical light (L) chains disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C; Superfamily: immunoglobulin C region; immunoglobulin homology  
C; Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin F; 20-85/Domain: immunoglobulin homology <IM1>  
F; 133-202/Domain: immunoglobulin homology <IM2>  
F; 239-306/Domain: immunoglobulin homology <IM3>  
F; 14/Disulfide bonds: interchain (to light chain) #status experimental  
F; 27-83, 140-200, 246-304/Disulfide bonds: #status experimental  
F; 102, 103, 106, 109/Disulfide bonds: interchain (to heavy chain) #status experimental  
F; 176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 51.3%; Score 1148.5; DB 1; Length 326;  
Best Local Similarity 90.3%; Pred. No. 1.1e-63;  
Matches 214; Conservative 8; Mismatches 8; Indels 7; Gaps 2;

QY 170 STOKVDT-----HTCPPCPAPDLGGPSVFLEPPPKDFTLMISRTPEVTCVVVDVSHE 223  
| | | | | : : : : :  
Db 90 SNTKVDTKVERKCCEPCPPAPP-VAGSPSVFLPFPKPDKTLMISRTPEVTCVVVDVSHE 148  
QY 224 DPVKENWYVDGEVHNATKPREEOYNSTPRVSVLTVLHODWLNGKEYCKVSNKALP 283  
| | | | | : : : : :  
Db 149 DPEQENWYVDGEVHNATKPREEFNFNTRVSVLTVLVHQDWLNKKEYCKVSNKGLP 208  
QY 284 APIEKTISKAKGPREFPVVTLPSPSRDELTKNOVSLTCLVKGYPSDIAVEWESNQPEN 343  
| | | | | : : : : :  
Db 209 APIEKTISKKGPREPVVTLPSPREEMTKNOVSLTCLVKGYPSDIAVEWESNQPEN 268  
QY 344 NYKTPPLVSDSGSFPLYSKLTVDKSRWQGNVFCFSVMHEALHNHYTKSLSLSPG 400  
| | | | | : : : : :  
Db 269 NYKTPPLMLSDSGSFPLYSKLTVDKSRWQGNVFCFSVMHEALHNHYTKSLSLSPG 325

RESULT 9  
G4HU  
Ig gamma-4 chain C region - human  
C; Species: Homo sapiens (man)  
C; Date: 02-Apr-1982 #sequence revision 02-Apr-1982 #text\_change 16-Jul-1999  
C; Accession: A90933; A90249; A02150  
R; Ellison, J.; Buxbaum, J.; Hood, L.  
DNA 1, 11-18, 1981  
A; Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.  
A; Reference number: A90933; MUID: 83157104  
A; Accession: A90933



A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I haplot  
A:Reference number: A91749; MUID:84030930  
A:Accession: A91749  
A:Molecule type: mRNA  
A:Residues: 1-323 <BER>  
A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-Thr  
R:Pratt, D.M.; Mole, L.E.  
Biochem. J. 151, 337-349, 1975  
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunoglobulin  
A:Reference number: A90290; MUID:76135469  
A:Accession: A90290  
A:Molecule type: protein  
A:Residues: 1-47, 'E', '49-71', 'PV', '72-128 <PRA>  
R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
A:Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy chain  
A:Reference number: A93928; MUID:83299917  
A:Accession: A93928  
A:Molecule type: mRNA  
A:Residues: 88-103, 'M', '105-143, 'E', '145-184, 'A', '186, 'E', '188-266 <MAR>  
A:Cross-references: GB:M16426; MUID:9165111; PIDN:AAA31289.1; PID:9165112  
A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic marker  
R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.  
Biochem. J. 116, 249-259, 1970  
A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin  
A:Reference number: A90245; MUID:70110015  
A:Accession: A90245  
A:Molecule type: protein  
A:Residues: 132-143, 'E', '145-161 <FRU>  
R:Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
In Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiksell.  
A:Reference number: A94416  
A:Accession: A94416  
A:Molecule type: protein  
A:Residues: 129-131;155-172, 'D', '174-184, 'A', '186, 'E', '188-200, 'D', '202-217, 'E', '219-232, 'Q',  
A:Note: this has the e15 allotypic marker, 185-Ala  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap)  
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-82/Domain: immunoglobulin homology <IM1>  
F:130-199/Domain: immunoglobulin homology <IM2>  
F:236-303/Domain: immunoglobulin homology <IM3>  
F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.2%; Score 923; DB 1; Length 323;  
Best Local Similarity 61.5%; Pred. No. 7.5e-50;  
Matches 177; Conservative 33; Mismatches 46; Indels 32; Gaps 5;  
Qy 123 RCPDGFSS-----NETSSKAPCKRHTNCSVFGLLLLTQKGNATHDNCISGSESTOKVDKT 177  
Db 57 ROSSGLYSLSSVSVTSQQ-----VTCNV-----AHPATNTKVDKT 94  
Qy 178 ---HFC---PPCAPPELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFNW 232  
Db 95 VAPSTCKPTCPPELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSDQDPEVQFTWY 154  
Qy 233 VDGVEVHNATKPREQNSYRVVSVLTVLHODWLNKGVCKVSKNKPAPTEKTISK 292  
Db 155 INEQVARTPRPLRQQNSIRVSVLTPIHQDWLKGKFKCKVHNKALPAPTEKTISK 214  
Qy 293 AKGQPREQVYTLPPSRDELTKNOVSLCLVKGFPSDIAVEWESNGQPENNYKTPPV 352  
Db 215 AKGQPLEPKVYTMGPPELSRSVSLTCMNGFPSPDISVEWENKGAEDNYKTPPAVL 274  
Qy 353 DSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 400  
Db 275 DSDGSFFLYNKLVSPTSEWQGDVFTCSVMHEALHNHYTQKSISRSPG 322

RESULT 12  
147159

Ig gamma 2b chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47159  
R:Kacskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of  
A:Reference number: I47158; MUID:95015845  
A:Accession: I47160  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03780; MUID:9433125; PIDN:AAA52218.1; PID:9433126  
C:Genetics:  
A:Gene: IgG2b  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 41.0%; Score 917.5; DB 2; Length 328;  
Best Local Similarity 60.5%; Pred. No. 1.7e-49;  
Matches 176; Conservative 35; Mismatches 43; Indels 37; Gaps 6;  
Qy 125 PDGFFSNETSSKAPC-----RKHTNCSVFGLLLLTQKGNATHDNCISGSESTQKVDK---T 177  
Db 59 PSGLYSLSSVMVTPASSLSKSYTCNV-----NHPATTTKVDKRVGT 100  
Qy 178 HTCPPCP-----APELLGGPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVKFNW 231  
Db 101 KTKPPCPCIPACESP-----GPSVFIFPPPKPDTLMISRTPEVTCVVVDVSDQENPEVQFSW 156  
Qy 232 YVDGVEVHNATKPREQNSYRVVSVLTVLHODWLNKGVCKVSKNKPAPTEKTIS 291  
Db 157 YVDGVEVHTATFRKEEGFNSTYRVSVLPIHQDWLNGKFKCKVNNKNDLPAPETRIIS 216  
Qy 292 KAKGQPREQVYTLPPSRDELTKNOVSLCLVKGFPSDIAVEWESNGQ---PENNYKTPP 349  
Db 217 KAKGQPREQVYTLPPHAEELSRKSVITCLVGFYPPDIDVWQNGOPEPEGNRYRTP 276  
Qy 350 PVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 400  
Db 277 PQDDVDGTYFLYKFSVDKASWQGGIFQCAVMHEALHNHYTQKSISKTPG 327

RESULT 13  
147159  
Ig gamma 2a chain constant region - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
C:Accession: I47159  
R:Kacskovics, I.; Sun, J.; Butler, J.E.  
J. Immunol. 153, 3565-3573, 1994  
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of  
A:Reference number: I47158; MUID:95015845  
A:Accession: I47159  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-328 <KAC>  
A:Cross-references: EMBL:U03779; MUID:9433123; PIDN:AAA52217.1; PID:9433124  
C:Genetics:  
A:Gene: IgG2a  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 41.0%; Score 917.5; DB 2; Length 328;  
Best Local Similarity 60.5%; Pred. No. 1.7e-49;  
Matches 176; Conservative 35; Mismatches 43; Indels 37; Gaps 6;  
Qy 125 PDGFFSNETSSKAPC-----RKHTNCSVFGLLLLTQKGNATHDNCISGSESTQKVDK---T 177  
Db 59 PSGLYSLSSVMVTPASSLSKSYTCNV-----NHPATTTKVDKRVGT 100





Thu Mar 1 09:58:04 2001

us-09-389-782a-5.rpr

Page 8

Db 283 ADGSYFLYSKLTVDKSAWDQGTVTCTSVNHEALHNVHTOKAISRSPG 329

Search completed: March 1, 2001, 09:15:46  
Job time: 144 sec

PI Laufter L, Oquendo P, Zett



```

Db 368 nstyrsvsvltvlgdwlngdyckvsnkalpampqktiskakgqprepyvtlppsrd 427
QY 311 ELTKNOVSLTCLVKGYPSPDIAVWESNGOPENNYKTTTPVLDSDGSFELYSKLTVDKSR 370
Db 428 eltkngvslclvkgfyrhriavwesngppennnykttppvldsdgsfflyskltvdksr 487
QY 371 WQGNVFCSCVMHEALHNYTKSLSPG 400
Db 488 wqgnvfscvmhealhnhytqkslspsg 517

RESULT 3
ID W60037 standard; Protein; 376 AA.
XX
AC W60037;
XX
DT 11-SEP-1998 (first entry)
XX
DE Antigenic peptide hFas (nd29) containing Fc region.
XX
KW Fas ligand; Fas antagonist; apoptosis related disease; liver disease;
KW heart failure; kidney failure; graft-versus-host disease; antibody;
KW myocardial infarction; ischemic restenosis; endotoxemic shock.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT Protein
FT /note= "hFas antigen signal peptide"
FT /note= "hFas (nd29) protein"
XX
PN W09818487-Al.
XX
PD 07-MAY-1998.
XX
PF 31-OCT-1997; 97WO-JP03978.
XX
PR 26-SEP-1997; 97JP-0262521.
PR 31-OCT-1996; 96JP-0290459.
PR 27-DEC-1996; 96JP-0351718.
XX
PA (MOCH ) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSAKA BIOSCIENCE INST.
XX
PI Nagata S, Suda T, Yatomi T;
XX
DR WPI; 1998-271925/24.
DR N-PSDB; V34430.
XX
PT Use of Fas antagonist for treatment and prevention of
PT apoptosis-related diseases - such as heart or kidney failure,
PT graft-versus-host disease or liver disease
XX
PS Examples; Fig 5-9; 86pp; Japanese.
XX
CC This represents the antigenic peptide hFas (nd29) containing the Fc
CC region. The invention provides the use of Fas antagonist as an agent for
CC the treatment and prevention of apoptosis-related diseases. The Fas
CC antagonist can be a partial Fas antigen peptide containing the
CC extracellular part of the protein, but lacking the signal sequence, an
CC anti-Fas antibody, or an anti-Fas ligand antibody, where the antibody is
CC preferably a humanised antibody. The Fas antagonist is used in the
CC treatment and prevention of diseases such as myocardial infarction, heart
CC failure, ischemic heart disease, acute kidney failure, graft-versus-host
CC disease, ischemic restenosis of the heart, liver or kidney, and
CC endotoxemic shock, and also as an organ preservative in transplantation.
CC The agent is of low toxicity but effectively inhibits the Fas/Fas ligand
CC system.
XX
SQ Sequence 376 AA;

```

```

Query Match 60.4%; Score 1352; DB 19; Length 376;
Best Local Similarity 67.3%; Pred No. 5e-78;
Matches 270; Conservative 15; Mismatches 60; Indels 56; Gaps 9;

QY 8 LHYDEFTSHQLCDCKPCPGTYLKQHCCTAKW-KTVCAPCPD-HYYTDSWHITSDCLYCSVP 65
Db 23 lhhdgqfchk-pppggerkardctvngdepdcvqdegkeytdkahfsskrrcr-l 77
QY 66 CKELQYVQOE--CNRTHNRVCECKEGRYLEIEFCLKHRSCPP-----GFGVVQAGTPERNT 119
Db 78 cdeghgleveinctrtgntkrcrknfnstvc---ehcdpctkcehglikeitstnt 134
QY 120 VCKRCPDGFNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCISGNSESTQKVOKTHT 179
Db 135 kcke--egrsrsnepks-----cdktht 154
QY 180 CPPCPAPELIGFSVFLFPKPKDMLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVVH 239
Db 155 cpcpapellggsvflfppkpkdtlmsrtpevtcvvdvshedpevkfnwyvdgveh 214
QY 240 NAKTKPREQYNSTYRVSVLTVLHQDLNGKEYKCVSNKALPAIEKTIKAKGP 299
Db 215 naktkpreeqynstyrsvsvltvlgdwlngkeykcvsnkalpapiektiskakgpre 274
QY 300 PQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGOPENNYKTTTPVLDSDGSFF 359
Db 275 pqvyltppsrdeltkngvslclvkgfypsdiavwesngppennnykttppvldsdgsff 334
QY 360 LYSKLTVDKSRWQGNVFCSCVMHEALHNYTKSLSPG 400
Db 335 lyskltvdksrwqgnvfscvmhealhnhytqkslspsg 375

RESULT 4
W50287
ID W50287 standard; Protein; 376 AA.
XX
AC W50287;
XX
DT 16-JUL-1998 (first entry)
XX
DE Human Fas antigen derivative/IgG1 Fc fusion.
XX
KW Human; Fas antigen; derivative; apoptosis regulation; gene therapy;
KW treatment; diabetes; arthritis; lupus; hepatitis; Influenza; HIV;
KW apoptosis modulation; immunoglobulin G1 Fc; IgG1 Fc; fusion.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT /label= sig_peptide
FT /label= mat_peptide
XX
PN W09742319-Al.
XX
PD 13-NOV-1997.
XX
PF 01-MAY-1997; 97WO-JP01502.
XX
PR 02-MAY-1996; 96JP-0135760.
XX
PA (MOCH ) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSAKA BIOSCIENCE INST.
XX
PI Nagata S, Nakamura N;
XX
DR WPI; 1997-558981/51.
DR N-PSDB; V07004.
XX

```

PT Fas antigen derivative containing modified extracellular region -  
PT has low antigenicity, promotes apoptosis and is useful in treatment  
XX of viral and other diseases

PS Disclosure; Fig 4; 102pp; Japanese.

XX The present sequence is a Fas antigen derivative/IgG1 Fc  
CC fusion, which contains a Fas antigen extracellular region lacking  
CC one or more amino acid residues in the region from the  
CC amino-terminal to (but excluding) the 1st cysteine residue  
CC (preferably at least 29 residues are deleted).

CC The derivative is an effective regulator of apoptosis and can be  
CC used (either by administration of the polypeptide, or by the use  
CC of the coding DNA in gene therapy) to treat a range of diseases,  
CC e.g. diabetes, arthritis, lupus and in particular viral diseases  
CC such as hepatitis, influenza and HIV, by modulating apoptosis of  
CC virus-infected cells.

XX Sequence 376 AA;

Query Match 59.78; Score 1337; DB 18; Length 376;  
Best Local Similarity 66.68; Pred. No. 4.8e-77;  
Matches 267; Conservative 16; Mismatches 62; Indels 56; Gaps 9;

QY 8 LHYDEETSHQLLCKDPCPGTYLKQHTAKW-KTVCAPCPD-HYTYDSWHTSDECLYCSVPV 65

DB 23 lhhdgqfchk-----pcppgkardctvngdepdcvpcqeqkeytdkahfskrrrr-i 77

QY 66 CKELQYVQKE--CNRTNHRVCECKEGRYLETEFLKHKRSCPP-----GFGVVQAGTPEPNT 119

DB 78 cdeghgieveincrtntqtkercpknfcnstvc---ehcdpctkcehgliiecktltsnt 134

QY 120 VCKRCPDGFSNETHSSKAPCKRHTNCVFGLLLTQKGNATHDNCNSSESTQKVDKTHT 179

DB 135 kcke-egrsneks-----cdktht 154

QY 180 CPPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 239

DB 155 cpcpapelggpsvflfpkpkdtlmsrtpevtcvvdvshedpevkfnwvydgvvsh 214

QY 240 NAKTKPREEQYNSTYRVVSVLTCLVKGFPSPDIAVWESNGOPENNYKTTTPVLDSDGSEF 299

DB 215 naktkpreeqynstyrsvsvltvqhgdwngkeykckvsnkalpapietkiskagqpre 274

QY 300 POVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGOPENNYKTTTPVLDSDGSEF 359

DB 275 pqvylppsrdeltknqvsltcvkgfypsdiavewesngqpnykttppvldsdgsff 334

QY 360 LYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPG 400

DB 335 lysnltvdksrwqgnvfscsvmhealhnhytqkslsisp 375

RESULT 5

R81882

ID R81882 standard; Protein; 438 AA.

XX AC

XX R81882;

XX 30-MAR-1996 (first entry)

XX Plasmid pDC406/OX40/Fc\* encoding an OX40/Fc mutein protein.

XX OX40: OX40-L; cytokine; cell surface molecule; plasmid;

XX pDC406/OX40/Fc\*; membrane glycoprotein.

XX Synthetic.

XX OS

XX US5457035-A.

XX PN

XX 10-OCT-1995.

XX

PF 23-JUL-1993; 93US-0097827.

XX PR 23-JUL-1993; 93US-0097827.

XX (IMMV ) IMMUNEX CORP.

XX Baum PR, Fanslow WC, Gayle RB, Goodwin RG;

XX WPI; 1995-357992/46.

XX N-PSDB; T00829.

XX New isolated DNA encoding the OX40 ligand polypeptide - also vectors

XX and host cells, used to produce recombinant ligand used in e.g.

XX prim. T cell culture, to modulate immune response etc.

XX Example 2; Column 35-38; 26pp; English.

XX This plasmid encodes an OX40/Fc antibody fragment mutein protein,

XX and is used to express a soluble OX40/Fc mutein fusion protein for

XX use in detecting cDNA clones encoding a OX40 ligand. The Fc

XX fragment may be derived from human IgG1, and the plasmid may be

XX used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell

XX line. Culture supernatant was purified by affinity chromatography

XX and this was used, together with labeled goat anti-human IgG to

XX screen various cell lines.

XX Sequence 438 AA;

QY Query Match 59.5%; Score 1332; DB 16; Length 438;

DB Best Local Similarity 62.4%; Pred. No. 1.2e-76;

Matches 262; Conservative 23; Mismatches 75; Indels 60; Gaps 6;

QY 20 CDKCPPTTYLKQHTAKWKTVCAPCPDHPHYTD--SWHTSDECLYCSVCKELQVVKQECN 77

DB 39 crecqpghgmvrchdtrdtlchpcetgfyneavnydtkcqtqcnh--rsgsalkqnt 96

QY 78 RTHNRVCECKEGRYLETEFLKHKRSCPPGFGVQAGTPERNVCKRCPCDGFSSNETSSA 137

DB 97 ptqdtvcr-----pgtqprqdsykyigdvcpccpghfs--pgnng 137

QY 138 PCRKHTNCVFGLLLTQKGNATHDNCNSGNS-----EST----- 171

DB 138 ackpwtntclsgkqrhpsasdlavcedrslatlletqrptfrptvtqstevwpts 197

QY 172 -----QKVDKHTTCTPCPCAPPELLGSPSVFLFPKPKDTLMISRTPEVTCVVVDV 220

DB 198 elpstptliveprscdkthtccpccapeagapsvflfpkpkdtlmsrtpevtcvvvdv 257

QY 221 SHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTCLVKGFPSPDIAVWESNGQ 280

DB 258 shedpevkfnwvydgvvshvhnaktpreeqynstyrsvsvltvqhgdwngkeykckvsnk 317

QY 281 ALPAPIEKTISKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNGQ 340

DB 318 alpapietkiskagqprepvytlppsrdeltknqvsltcvkgfypsdiavewesngq 377

QY 341 PENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPG 400

DB 378 pennykttppvldsdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsisp 437

RESULT 6

W48976

ID W48976 standard; Protein; 438 AA.

XX AC

XX W48976;

XX 25-SEP-1998 (first entry)

XX OX40/Fc mutein.

XX OX40/Fc; cytokine; T cell antigen; TH-2 immune response; OX40-L;

XX KW

chimeric.  
 Chimeric - Homo sapiens.  
 Chimeric - Mus sp.

Key Location/Qualifiers  
 Region 1..206  
 /note= "Extracellular domain of mouse OX40"  
 Region 207..438  
 /note= "Mutant Fc region of human IgG1 antibody"  
 Misc-difference 225  
 /note= "changed from Leu in wild-type to Ala in mutant"  
 Misc-difference 226  
 /note= "changed from Leu in wild-type to Gly in mutant"  
 Misc-difference 228  
 /note= "changed from Gly in wild-type to Ala in mutant"

US5783665-A.  
 21-JUL-1998.  
 22-JUN-1995; 95US-0494574.  
 23-JUL-1993; 93US-0097827.  
 22-JUN-1995; 95US-0494574.  
 (IMMUNEX CORP.)

Baum PR, Fanslow WC, Gayle RB, Goodwin RG;  
 WPI: 1998-427099/36.  
 N-PSDB; V32636.

Purified polypeptide OX-40 ligands - for co-stimulation of T-cell production and binding assays for OX-40 and homologues  
 Example 2; Col 37-40; 26pp; English.

The present sequence represents the OX40/Fc fusion protein that contains the extracellular domain of mouse OX40 fused to the mutated Fc region of the human IgG1 antibody. The fusion protein was used for detecting cDNA clones encoding an OX40 ligand. The invention claims for a murine OX40-L cytokine (W48975) that binds to the murine T cell antigen, OX40. The OX40-L protein is claimed to be useful for co-stimulation of T-cell production and in binding assays for detecting OX40 or its homologues. The OX40-L protein is also claimed to generate a TH-2 immune response.

Sequence 438 AA;

Query Match 59.5%; Score 1332; DB 19; Length 438;  
 Best Local Similarity 62.4%; Pred. No. 1.2e-76;  
 Matches 262; Conservative 23; Mismatches 75; Indels 60; Gaps 6;

20 CDKCPGTYLKQCTAKWTCVAPCPDPHYTD--SWHTSDCLYCSVPVKELQYVKQECN 77  
 39 cecqpggmvrncdtrdtlchpcetgfyneavnydtkcqtqcnh--rsgselkqnc 96  
 78 RTHNRYCECKEGRYLEFCLKHRSPPCGVVOAGTPERTVCKRCPDGFNFSSKA 137  
 97 ptdgtvcr-----pgtqrqdsqgkylgdcvpcpghfs--pgnnq 137  
 138 PCRKHTNCSVFGLLLRQKGNATHDNICSGNS-----EST----- 171  
 138 ackpwnctslgkqrhpsadsladvedrslatlilwetqrftptvtqstvtwprts 197  
 172 -----QKVDKTHTCPPCAPPELLGGPSVFLFPKPKDLMISRTPEVTCVVDV 220  
 198 elpstpvlveprscdkthtccpcaeaegapsvflfpkpkdltlmsrtpevtcvvdv 257

QY 221 SHEDPEVKFNWYDGVGVHNAKTPREOYNSTYRVVSVLTVLHQDWLNGKVKCKVSNK 280  
 DB 258 shedpevkfnwYDGVGVHNAKTPREOYNSTYRVVSVLTVLHQDWLNGKVKCKVSNK 317  
 QY 281 ALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQ 340  
 DB 318 alpapiEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQ 377  
 QY 341 PENNYKTPPVLDSGDSFFLYSKLTVDKSRWQQGNVPSFCVSMHEALHNHTYTKSLSPG 400  
 DB 378 pennyKTPPVLDSGDSFFLYSKLTVDKSRWQQGNVPSFCVSMHEALHNHTYTKSLSPG 437

RESULT 7  
 Y68949  
 ID Y68949 standard; Protein; 764 AA.  
 XX  
 AC Y68949;  
 DT  
 XX 30-MAY-2000 (first entry)  
 DE Fusion protein of murine delta-related protein and human IgG Fc.  
 XX  
 KW Cell development cycle; Delta family; membrane surface-bound ligand;  
 KW endothelial cell biology; gene therapy; subcortical infarct;  
 KW cerebral autosomal dominant arteriopathy; leucoencephalopathy;  
 KW ischemic stroke; chimera.  
 XX  
 OS Chimeric - Mus sp.  
 OS Chimeric - Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Protein 1..529  
 /note= "extracellular region of the murine  
 FT Protein 533..764  
 /note= "human IgG Fc portion"  
 FT  
 XX WC2000006726-A2.  
 XX  
 PD 10-FEB-2000.  
 XX  
 PF 12-JUL-1999; 99WO-US15710.  
 XX  
 PR 27-JUL-1998; 98US-0123168.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Shutter JR, Stark KL;  
 XX  
 DR WPI; 2000-195294/17.  
 DR N-PSDB; Z60926.  
 XX  
 XX Cell development cycle protein of delta family useful for treating  
 PT various disorders associated with central nervous system e.g. cerebral  
 PT autosomal dominant arteriopathy and ischemic strokes  
 XX  
 PS Example 6; Page 169-171; 171pp; English.  
 XX  
 CC The present sequence represents a fusion protein of the extracellular  
 CC domain of a murine polypeptide, which is a member of the cell  
 CC development cycle protein family known as the Delta family of  
 CC mammalian membrane surface-bound ligands, and the human immunoglobulin  
 CC (IgG) Fc portion. The murine delta-related protein gene is expressed  
 CC within vascular endothelium indicates a role for the polypeptides  
 CC in the control of endothelial cell biology. The murine polynucleotide  
 CC is identified from a white adipose tissue cDNA library. The polypeptide  
 CC is useful for identifying receptors, which bind to and/or are activated  
 CC by the polypeptide. The polynucleotide is useful in gene therapy of  
 CC cerebral autosomal dominant arteriopathy with subcortical infarcts and  
 CC leucoencephalopathy, an autosomal dominant disorder causing ischemic  
 CC strokes.





```

RESULT 9
W73513
ID W73513 standard; Protein; 388 AA.
XX
AC W73513;
XX
DT 02-MAR-1999 (first entry)
XX
DE Rabbit TGFbetaRII:Fc protein.
XX
DE Transforming growth factor-beta receptor; TGF-beta receptor; arthritis;
XX KW fusion protein; fibroproliferative disorder; diabetic nephropathy;
XX KW glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;
XX KW collagen vascular disorder; therapy; rabbit.
XX
OS Oryctolagus sp.
XX
XX W09848024-A1.
XX
XX 29-OCT-1998.
XX
XX 16-APR-1998; 98WO-US07587.
XX
XX 18-APR-1997; 97US-0044641.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Cate R, Gotwals P, Koteliarsky V, Sanicola-Nadel M;
XX
XX WPI; 1998-609994/51.
XX
XX N-PSDB; V08998.
XX
XX Transforming growth factor-beta receptor fusion protein - used to
XX treat fibroproliferative disorders
XX
XX Claim 4; Page 18-19; 70pp; English.
XX
XX This sequence is a fusion protein of the rabbit transforming growth
XX factor-beta receptor II (TGF-betaRII) and an antibody Fc region. The
XX encoded protein is an example of a protein of the invention, which are
XX isolated TGF-beta receptor fusion proteins that competitively inhibit
XX binding of TGF-beta to TGF-beta receptor. The fusion protein can be used
XX in a method for lowering the levels of TGF-beta in an individual having
XX arthritis. It can also be used to treat medical conditions such as
XX fibroproliferative disorders. The fibroproliferative disorder is a
XX kidney, intraocular or pulmonary fibrosis, especially selected from
XX diabetic nephropathy, glomerulonephritis, proliferative
XX vitreoretinopathy, or myelofibrosis. The fusion proteins can also be used
XX to treat collagen vascular disorders, including systemic sclerosis,
XX polymyositis, scleroderma, dermatomyositis, and systemic lupus
XX erythematosus. They can also be administered following coronary
XX angioplasty, to prevent restenosis or scarring and reclosing of arteries.
XX
XX Sequence 388 AA;

Query Match 55.9%; Score 1252.5; DB 19; Length 388;
Best Local Similarity 65.7%; Pred. NO. 1e-71;
Matches 257; Conservative 17; Mismatches 70; Indels 47; Gaps 10;

QY 37 WKTVCAPCPDHYH---TDSWHTSDE-----CLYCV---SPVKRELQVVKQECNRTHN 81
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 17 wtriastipbhvdkvsnndmmvtdngavkfpqlckfdvrsctdnqkscmncsit-- 74
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 82 RVCECKEGRYLEIFCLKLRSCPPGFVQVQAGTPERTNVCKRCPD-----GFFSNETSS 135
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 75 sice-----kahevcv-----avvrkndenitltvchdpxklayhgflledsas 118
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 136 KAPCRKHTNCSVFG---LLLTQGNATHDNCISGNSESTQK---VKTHTCPCPCPAPELL 189
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 pkcimkek--kvifgetffmcsctdcndhifseeyttsspdldvdkthctcpcpapell 176
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

190 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKNNWYVDGVEVHNATKPREQ 249  
 177 gpsvflfppkpkdtlmisrtpevtcvvdvshedpevkfnwyvdgvevhnaktckpreeq 236;  
 250 YNSTYRVSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKAKGQRPQVYTLPPSR 309  
 237 ynstyrsvsvltvlhqdwlngkeykckvsnkalpapiektiskakgqrpqvylppsr 296  
 310 DELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPVLDSDGSFFLYSKLTVDKS 369  
 297 deltknqvslclvkgfypsdiavewesngqpennnykttppvidsdgsfflyskltvdk 356  
 370 RWQGNVFSVSMHEALHNHYTQKSLSLSPG 400  
 357 rwqgnvfscsvmhealhnhytqkslspsg 387

RESULT 10  
 Y54063  
 ID Y54063 standard; Protein; 388 AA.  
 XX  
 AC Y54063;  
 XX  
 DT 27-MAR-2000 (first entry)  
 XX  
 DE Amino acid sequence of TGF-beta type II receptor variant/IgG1 fusion.  
 XX  
 KW Splice variant; rabbit; transforming growth factor-beta; TGF-beta;  
 KW type II receptor; Fc portion; human IgG1; fusion protein; arthritis;  
 KW fibroproliferative disease; renal; intra-ocular; pulmonary; fibrosis;  
 KW diabetic nephropathy; glomerulonephritis; collagen vascular disease;  
 KW proliferative vitreoretinopathy; myelofibrosis; systemic sclerosis;  
 KW polymyositis; scleroderma; dermatomyositis; systemic lupus erythematosus;  
 KW restenosis; wound; connective tissue production; adhesion; scarring;  
 KW post-radiation fibrosis.  
 XX  
 OS Synthetic.  
 OS Oryctolagus cuniculus.  
 OS Homo sapiens.  
 XX  
 PN W09965948-A1.  
 XX  
 PD 23-DEC-1999.  
 XX  
 PF 16-JUN-1999; 99WO-US13629.  
 XX  
 PR 16-JUN-1998; 98US-0089452.  
 XX  
 XX (BIOJ ) BIOGEN INC.  
 XX  
 XX Koteliarsky V, Gotwals P, Cate R, Sanicola-Nadel M;  
 DR WPI; 2000-106083/09.  
 DR N-PSDB; Z45251, Z45252.  
 XX  
 XX New fusion protein of a splice variant of transforming growth  
 PT factor-beta receptor, for inhibiting the growth factor, e.g. in  
 PT treatment of fibrosis  
 XX  
 PS Disclosure; Page 62-63; 69pp; English.  
 XX  
 CC The present sequence represents a splice variant of a rabbit transforming  
 CC growth factor-beta (TGF-beta) type II receptor fused to the Fc portion  
 CC of human IgG1. The fusion protein has higher affinity for TGF-beta than  
 CC fusion proteins comprising the non-variant form of the receptor. The  
 CC fusion protein contains soluble TGF-beta receptor constructs that are  
 CC devoid of a transmembrane region (and are secreted from the cell) but  
 CC retain the ability to bind TGF-beta. The protein competitively inhibits  
 CC binding of TGF-beta to cellular receptors and/or forms an inactive  
 CC complex with TGF-beta. The protein is used to reduce levels of TGF-beta,  
 CC for treatment of arthritic conditions associated with overexpression  
 CC of TGF-beta, especially fibroproliferative diseases, e.g. renal,  
 CC intra-ocular or pulmonary fibrosis; diabetic nephropathy;

CC glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;  
 CC collagen vascular disease, e.g. systemic sclerosis, polymyositis,  
 CC scleroderma, dermatomyositis or systemic lupus erythematosus; and  
 CC fibrosis associated with restenosis. It is also used for treating  
 CC wounds, to prevent overproduction of connective tissue and so prevent  
 CC adhesions or scarring, and to prevent post-radiation fibrosis (by  
 CC administration to patients about to undergo radiation therapy).  
 XX  
 SQ Sequence 388 AA;

Query Match 55.9%; Score 1252.5; DB 21; Length 388;  
 Best Local Similarity 65.7%; Pred. No. 1e-71;  
 Matches 257; Conservative 17; Mismatches 70; Indels 47; Gaps 10;

QY 37 WKTVCAPCPDHY---TDSWHTSDE-----CLYC---SPVCKELQYVVKQECNRTHN 81  
 Db 17 wtrlastipphvqsvnmndmvtndngavkfpqickfcdvrsctcdngkscmsncsit-- 74  
 QY 82 RVCCKEGRYLEIEFCLKHSRCPGFGVVQAGTPERNVCKRCPP-----GFFSNETSS 135  
 Db 75 sice-----kahevcv---avwrkndenitltvchdpklayhgfiledsas 118  
 QY 136 KAPCRKHTNCSVFG---LLQKGNATHDNTCSGSESTQK---VDKTHTCPPCPAPPELL 189  
 Db 119 pkcimkek--kvfgetffmcscstdecdndhiifseeeyttspdlvdkthtccppcpapell 176  
 QY 190 GGPSVFLPPKPKDTLMSRPETVTCVVVDYSHEDPEVKFNWYDGVVHNAKTPREEQ 249  
 Db 177 ggsavilfpkpkdtlmsrtpetvctcvvdvshedpevkfnwvdygvevhnaaktpreeq 236  
 QY 250 YNSYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTIKSKAGQPREQVYTLPPSR 309  
 Db 237 ynstyrsvsvltvlhqdwlngkeyckvsnkalpapiektiskakgpreqvyltlpsr 296  
 QY 310 DELTKNQSLSLCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKKS 369  
 Db 297 deltknqvsllclvkgyfypsdiawesngqpennnykttppvldsdgsfflyskltvdk 356  
 QY 370 RWQGNVFCSCVMHEALHNHYTQKSLSLSPG 400  
 Db 357 rwqgnvfscsvmhealhnhytqkslsispg 387

RESULT 11  
 W71603  
 ID W71603 standard; Protein; 564 AA.  
 AC W71603;  
 XX  
 XX  
 XX 19-NOV-1998 (first entry)  
 DT  
 DE Human neuturin receptor alpha/Fc sequence (If2a) fusion protein.  
 XX  
 KW Human; neuturin receptor alpha; NTNR-alpha; variant; chimeric;  
 KW fusion protein; immunoadhesion; ret-expressing cell; neurological;  
 KW renal; haematological disease.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX W09836072-A1.  
 PN  
 XX  
 PD 20-AUG-1998.  
 XX  
 PF 17-FEB-1998; 98WO-US03179.  
 XX  
 XX 24-OCT-1997; 97US-0957063.  
 PR 18-FEB-1997; 97US-0802805.  
 PR 09-JUN-1997; 97US-0871913.  
 XX  
 XX (GETH ) GENENTECH INC.

PI Hynes MA, Klein RD, Rosenthal A;  
 XX  
 DR WPI: 1998-467175/40.  
 DR N-PSDB; V58006.  
 XX  
 PT New polypeptide(s) based on human neuturin receptor alpha and  
 PT related nucleic acid - useful for increasing survival of  
 PT ret-expressing cells for treating e.g. neurological, renal and  
 PT haematological diseases  
 XX  
 PS Example 3; Page 81-83; 116pp; English.

XX The present sequence represents human neuturin receptor alpha  
 CC (NTNR-alpha)/Fc sequence (If2a) fusion protein, from an example of the  
 CC present invention. NTNR-alpha proteins can be used: (a) to identify  
 CC molecules that bind specifically to it (potential agonists and  
 CC antagonists) and to purify such compounds; (b) to modulate response of  
 CC cells to neuturin (NTN); (c) to increase survival of Ret-expressing  
 CC cells or to activate Ret on the surface of cells (soluble glial derived  
 CC neurotrophic factor receptor (GDNFR alpha) may be used the same way);  
 CC (d) to increase the half-life of cognate ligands (especially NTN); (e)  
 CC diagnostically to determine serum levels of its ligands; and (f) as  
 CC animal feed additive or molecular weight marker. NTNR-alpha, its genes,  
 CC (antagonists and antisense nucleic acids, are useful in vivo or ex vivo  
 CC for treating conditions related to abnormal NTN activity or response,  
 CC particularly neurological (central or enteric), renal or haematopoietic  
 CC (spleen) diseases or injuries. Ab may be agonists or antagonists for  
 CC therapeutic use (e.g. as antagonists to treat excessive/unwanted NTNR-  
 CC alpha expression, e.g. in some tumours), also reagents for immunoassay  
 CC and affinity purification.  
 XX

SQ Sequence 664 AA;

Query Match 55.8%; Score 1249.5; DB 19; Length 664;  
 Best Local Similarity 63.6%; Pred. No. 3e-71;  
 Matches 264; Conservative 23; Mismatches 57; Indels 71; Gaps 14;

QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDECLYCSPVCKELQYV 72  
 Db 274 ncrasyqlvts-cpadnyqacglsyagmigfdmtpnyvdsptg---ivspwc----- 323  
 QY 73 KOECNRTHNRVCECKEGRYLEIEF---CLKH-----RSCP--PGFGVVQAGTP 115  
 Db 324 --scrgsgnmeece--kflr-dftenpclrnaigaqfngtdvnpkpsfqatqprv 378  
 QY 116 ERNVTVCRCPDGFFSNETSSKAPCRKHTNCSVFGLLLTQGNATHD-----NICS 165  
 Db 379 ekt---pslpddlsdstslgtsttctsvqeggl---kannskelsmcfelttnlip 431  
 QY 166 GNSESTQKVDKTHTCPPCPAPPELLGSPVFLFPKPKDTLMSRTPETVCVVVDVSHEDP 225  
 Db 432 gprd---pvdkthccppcpapellgspvflfpkpkdtlmsrtpetvcvvvdshedp 488  
 QY 226 EVKENWYVDGVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAP 285  
 Db 489 evkfnwyvdgvevhnaaktpreeqnystyrsvsvltvlhqdwlngkeyckvsnkalpap 548  
 QY 286 IEKTYISKAGQPREQVYTLPPSDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNY 345  
 Db 549 iektiskakgpreqvtytlppsrreemtknqvsitclvkgyfypsdiawesngqpenny 608  
 QY 346 KTTTPVLDSGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 400  
 Db 609 kttppvlidsdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispg 663

RESULT 12  
 Y80123  
 ID Y80123 standard; Protein; 664 AA.  
 XX  
 AC Y80123;  
 XX



QY 201 PKDTLMISRTPEVTCVVDVSHEDPEVFNMYVDGVEVHNNAKTKPREQYNSTYRVSVL 260  
 DB 245 pkdtlmisrtpevtcvvdvshedpevkfnwyvdgvevhnaktkpreeqynstyrsvsl 304  
 QY 261 TVLHODWLNGKEYCKVSNKALPAPIETISKAKGQPREPOVYTLPPSRDELTKNOVSIL 320  
 DB 305 tvlhodwlngkeyckvsnkalpapietiskakgqprepqvtytlppsrdeltnqvsil 364  
 QY 321 CLVKGFPSPDIWESNGQPNKYKTPPPVLDSDGSPFLYSLKLTVDKSRWQOGNVSFCS 380  
 DB 365 clvkgfypsdiwvesngqpennkyktpppvldsdgspfllyskltvdksrwqognvfcs 424  
 QY 381 VMHEALHNYTQKSLSPG 400  
 DB 425 vmhealhnytqkslspsg 444

## RESULT 14

W1604  
 ID W71604 standard; Protein; 664 AA.

XX AC W71604;

XX DT 19-NOV-1998 (first entry)

XX DE Rat neuturin receptor alpha/Fc sequence (Iff2a) fusion protein.

XX KW Human; neuturin receptor alpha; NTNR-alpha; variant; chimeric;  
 fusion protein; immunoadhesion; ret-expressing cell; neurological;  
 renal; haematological disease.

XX OS Synthetic.

XX OS Homo sapiens.

XX OS Rattus sp.

XX PN WO9836072-A1.

XX PD 20-AUG-1998.

XX PF 17-FEB-1998; 98WO-US031179.

XX PR 24-OCT-1997; 97US-0957063.

XX PR 18-FEB-1997; 97US-0802805.

XX PR 09-JUN-1997; 97US-0871913.

XX PA (GETH ) GENENTECH INC.

XX PI Hynes MA, Klein RD, Rosenthal A;

XX DR WPI; 1998-467175/40.

XX DR N-PSDB; V58007.

XX PT New polypeptide(s) based on human neuturin receptor alpha and  
 related nucleic acid - useful for increasing survival of  
 ret-expressing cells for treating e.g. neurological, renal and  
 haematological diseases

XX PS Example 3; Page 84-86; 116pp; English.

XX CC The present sequence represents rat neuturin receptor alpha  
 (NTNR-alpha)/Fc sequence (Iff2a) fusion protein, from an example of the  
 present invention. NTNR-alpha proteins can be used: (a) to identify  
 molecules that bind specifically to it (potential agonists and  
 antagonists) and to purify such compounds; (b) to modulate response of  
 cells to neuturin (NTN); (c) to increase survival of Ret-expressing  
 cells or to activate Ret on the surface of cells (soluble glial derived  
 neurotrophic factor receptor (GDNFR alpha) may be used the same way);  
 (d) to increase the half-life of cognate ligands (especially NTN); (e)  
 diagnostically to determine serum levels of its ligands; and (f) as  
 animal feed additive or molecular weight marker. NTNR-alpha, its genes,  
 (ant)agonists and antisense nucleic acids, are useful in vivo or ex vivo  
 for treating conditions related to abnormal NTN activity or response,  
 particularly neurological (central or enteric), renal or haematopoietic

CC (spleen) diseases or injuries. Ab may be agonists or antagonists for  
 therapeutic use (e.g. as antagonists to treat excessive/unwanted NTNR-  
 alpha expression, e.g. in some tumours), also reagents for immunosay  
 CC and affinity purification.

XX SQ Sequence 664 AA;

Query Match 55.7%; Score 1248.5; DB 19; Length 664;

Best Local Similarity 62.7%; Pred. No. 3.4e-71;

Matches 261; Conservative 21; Mismatches 61; Indels 73; Gaps 11;

QY 32 HCTAKWTKVACPCP-DHY- - - - -YTDSWHTSDECLYSPVCKELOVY 72  
 DB 274 ncrasyrtits-cpadnyqacclgsyagmigfdmtpnyvdsnptg--ivvspwc----- 323

QY 73 KQECNRTNHRVCEKE- - - - -GRYLEIEFCLKHRSCPPGFGVWQAGT 114

DB 324 --ncrgsgnmeeceekflrdfenpcrlnaiaqafngtdvmspkpslp----atqapr 377

QY 115 PERNTVCKRCPCDGFSSNETSSKAPCRKHTNCVSFLLLTQKGNATHD- - - - -NIC 164

DB 378 vekt---pslpddlsdstslgtswittctsiqegll-----kannskelsmofltttnis 430

QY 165 SGNSESTQKVDKTHTCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHED 224

DB 431 pg---srdpvdkthtccpcapellggpsvflfppkpkdtlmisrtpevtcvvdvshed 487

QY 225 PEVFNWTVDGVVHNNAKTKPREQYNSTYRVSVLTVLHODWLNGKEYCKVSNKALPA 284

DB 488 pevfnwvvdgvevhnaktkpreeqynstyrsvslvtvlhqdwlngkeyckvsnkalpa 547

QY 285 PIEKTISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWESNGQPN 344

DB 548 piektiskagqprepqvtytlppsrdektknoqslvtclvkgfypsdiwvesngqpenn 607

QY 345 YKTPPVLDSDGSPFLYSLKLTVDKSRWQOGNVSFCSVMHEALHNYTQKSLSPG 400

DB 608 yktpppvldsdgspfllyskltvdksrwqognvfscsvmhealhnytqkslspsg 663

## RESULT 15

Y80124

ID Y80124 standard; Protein; 664 AA.

XX AC Y80124;

XX DT 19-MAY-2000 (first entry)

XX DE Rat NTNR alpha and IgG fusion protein SEQ ID NO:18.

XX KW Human; neuturin receptor alpha; NTNR alpha; splenic haematopoiesis;  
 anaemia; thrombocytopaenia; hypoplasia; haemorrhage.

XX OS Rattus sp.

XX OS Synthetic.

XX PN US6025157-A.

XX PD 15-FEB-2000.

XX PF 24-OCT-1997; 97US-0957063.

XX PR 18-FEB-1997; 97US-0038839.

XX PR 09-JUN-1997; 97US-0049818.

XX PI (GETH ) GENENTECH INC.

XX PI Hynes MA, Rosenthal A, Klein RD;

XX DR WPI; 2000-181808/16.

XX DR N-PSDB; 291460.







Db 121 CKRCPDFFSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCSESTQKCGIDVTL 180

RESULT 2

ID 000300 PRELIMINARY; PRT; 401 AA.

AC 000300; 060236;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)

DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11B).

GN TNFRSF11B OR OPG OR OCIF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=KIDNEY;

RX MEDLINE=97262071; PubMed=9108485;

RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,

RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,

RA Derose M., Elliott R., Colombero A., Tan H.-L., Trill G., Sullivan J.,

RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,

RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,

RA Suggs S., Boyle W.J.;

RT "Osteoprotegerin: a novel secreted protein involved in the regulation

RT of bone density.";

RL Cell 89:309-319(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=LUNG FIBROBLAST;

RX MEDLINE=98151033; PubMed=9492069;

RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,

RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,

RA Tsuda E., Morinaga T., Higashio K.;

RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and

RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits

RT osteoclastogenesis in vitro.";

RL Endocrinology 139:1329-1337(1998).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=PLACENTA;

RX MEDLINE=98351569; PubMed=9688283;

RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;

RT "Cloning and characterization of the gene encoding human

RT osteoprotegerin/osteoclastogenesis-inhibitory factor.";

RL Eur. J. Biochem. 254:885-891(1998).

CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES

CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY

CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO

CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN

CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.

CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY

CC SIMILARITY).

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,

CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN

CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN

CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.

CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

DR EMBL; AB002146; BAA25910.1; -

DR EMBL; AB008822; BAA32076.1; -

DR EMBL; AB008821; BAA32076.1; JOINED.

DR EMBL; U94332; AAB53709.1; -

DR HSP; P25942; ICDP.

DR MIM; 602643; -

DR INTERPRO; IPR001368; -

DR PFAM; PF00020; TNFR\_C6; 3.

DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.

DR PROSITE; PS50050; TNFR\_NGFR\_2; 2.

DR PRODOM; PD000771; -; 1.

Glycoprotein; Repeat; Cytokine; Signal.

KW SIGNAL 1 21

FT CHAIN 22 401

FT DOMAIN 23 183

FT REPEAT 23 63

FT REPEAT 64 106

FT REPEAT 107 143

FT REPEAT 144 201

FT DOMAIN 306 365

FT DISULFID 41 54

FT DISULFID 44 62

FT DISULFID 65 80

FT DISULFID 83 97

FT DISULFID 87 105

FT DISULFID 118 142

FT DISULFID 145 160

FT CARBOHYD 98 98

FT CARBOHYD 152 152

FT CARBOHYD 165 165

FT CARBOHYD 178 178

FT CARBOHYD 289 289

FT CONFLICT 263 263

SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 45.9%; Score 1043; DB 4; Length 401;

Best Local Similarity 99.4%; Pred. No. 1.6e-83;

Matches 179; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPPKYLHYDEETSHQLCDKCPGTYLKQHCCTAKWKTVCAPCPDHYTDSWHTSDECL 60

DB 22 ETPPKYLHYDEETSHQLCDKCPGTYLKQHCCTAKWKTVCAPCPDHYTDSWHTSDECL 81

QY 61 YCSPVCKELQYVQECNTHNRVCECKEGRYLETFELCKHRSCPPGPGVQAGTPERTV 120

DB 82 YCSPVCKELQYVQECNTHNRVCECKEGRYLETFELCKHRSCPPGPGVQAGTPERTV 141

QY 121 CKRCPDFFSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCSESTQKCGIDVTV 180

DB 142 CKRCPDFFSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNCSESTQKCGIDVTL 201

RESULT 3

008712 PRELIMINARY; PRT; 401 AA.

ID 008712; 070202;

AC 008712; 070202;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)

DE (OCIF).

GN TNFRSF11B OR OPG.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALE/C; TISSUE=KIDNEY;

RX MEDLINE=97262071; PubMed=9108485;

RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,

RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,

RA Derose M., Elliott R., Colombero A., Tan H.-L., Trill G., Sullivan J.,

RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,

RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,

RA Suggs S., Boyle W.J.;

RT "Osteoprotegerin: a novel secreted protein involved in the regulation

RT of bone density.";

RL Cell 89:309-319(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=129/OLA, AND NIH SWISS;

RX MEDLINE=98382527; PubMed=9714833;



RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,  
RA Higashio K.;  
RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)  
RT gene and its expression in embryogenesis.";  
RL Gene 215:339-343(1998).  
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY  
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,  
CC BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND  
CC PLACENTA. NOT DETECTED IN SPLEEN.  
CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT  
CC DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY  
CC 15 TO DAY 17.  
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
DR EMBL; U94331; AAB53708.1; -.  
DR EMBL; AB013898; BAA28269.1; -.  
DR EMBL; AB013903; BAA33388.1; -.  
DR EMBL; AB013899; BAA33388.1; JOINED.  
DR EMBL; AB013900; BAA33388.1; JOINED.  
DR EMBL; AB013901; BAA33388.1; JOINED.  
DR EMBL; AB013902; BAA33388.1; JOINED.  
DR HSSP; P25942; ICDP.  
DR MGD; MGI:109587; Opg.  
DR INTERPRO; IPR000488; -.  
DR INTERPRO; IPR001368; -.  
DR PFAM; PF00020; TNFR\_C6; 3.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
DR PROSITE; PS50050; TNFR\_NGFR\_2; 2.  
DR PRODOM; PD000771; -. 1.  
KW Glycoprotein; Repeat; Cytokine; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 401 OSTEOPROTEGERIN.  
FT DOMAIN 23 201 4 X TNFR-CYS.  
FT REPEAT 23 63 TNFR-CYS 1.  
FT REPEAT 64 106 TNFR-CYS 2.  
FT REPEAT 107 143 TNFR-CYS 3.  
FT REPEAT 144 201 TNFR-CYS 4.  
FT DOMAIN 306 365 DEATH DOMAIN.  
FT DISULFID 41 54 BY SIMILARITY.  
FT DISULFID 44 62 BY SIMILARITY.  
FT DISULFID 65 80 BY SIMILARITY.  
FT DISULFID 83 97 BY SIMILARITY.  
FT DISULFID 87 105 BY SIMILARITY.  
FT DISULFID 118 142 BY SIMILARITY.  
FT DISULFID 145 160 BY SIMILARITY.  
FT CARBOHYD 98 98 POTENTIAL.  
FT CARBOHYD 165 165 POTENTIAL.  
FT CARBOHYD 178 178 POTENTIAL.  
FT CARBOHYD 289 289 POTENTIAL.  
FT VARIANT 138 138  
FT VARIANT 161 161 R -> P (IN STRAINS 129/OLA AND NIH  
FT VARIANT 165 165 SWISS).  
FT VARIANT 165 165 I -> R (IN STRAINS 129/OLA AND NIH  
FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH  
FT VARIANT 288 288 S -> A (IN STRAINS 129/OLA AND NIH  
FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH  
FT SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;

Query Match 40.6%; Score 922.5; DB 11; Length 401;  
Best Local Similarity 48.3%; Pred. No. 6e-73;  
Matches 196; Conservative 35; Mismatches 102; Indels 73; Gaps 9;  
QY 1 ETTPPKYLHYDEETSHQLLCKDCPPTGTLKQHCTAKWKTCAPCPDHYTDSWHTSDECL 60

Db 22 ETTPPKYLHYDEETSHQLLCKDCPPTGTLKQHCTAKWKTCAPCPDHYTDSWHTSDECL 81  
QY 61 YCSPVCKELQYVQKQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVVOAGTPERTV 120  
Db 82 YCSPVCKELQVQKQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVVOAGTPERTV 141  
QY 121 CKRCPDGFSTSSKAPCKRKHNTCSVFGLLLTQKNATHDNCISGNSSTQKCGIDVT 180  
Db 142 CKRCPDGFSTSSKAPCKRKHNTCSVFGLLLTQKNATHDNCISGNSSTQKCGIDVT- 200  
QY 181 DKTHTCPPCAPPELLGGPSVFLPPPKPDTLMISRTPEVTVVVDVSHEDPEYKFNMYD 240  
Db 201 -----LCEEAFFREAVPTK-----IIPNWSLVLD---SLPGTKVN--AE 235  
QY 241 GVEVHNAKTPREEQVNSTYRVVSVLTVLHQDWLNGKEY-----CKVSV----- 284  
Db 236 SVE-----RIKRHSSEOETQOLLKWKHQRDOEMVKLIQIDIDLCESVQRHLGHSNLT 291  
QY 285 -----NKALEPIEKTISKAKQPREPOVYTL-----PPSRDELTKNQVSLTC 327  
Db 292 TEQLLALMESLPQKKISPEIERTRKTSSEQLLKLKSLRWIKNGDQDQTLKGLMTALKH 351  
QY 328 LVKGFYPSDIAVWESNGOPENNYKTPPVVLDSDGSGFFLYSKLTVD 373  
Db 352 LKTSHEPKVT-----HSLRKTMRFLHSFTWRYLYQKLFLE 387  
RESULT 4  
O08727 PRELIMINARY; PRT; 401 AA.  
ID O08727  
AC O08727  
DT 01-JUL-1997 (Tremblrel. 04, Created)  
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)  
DE (OCIF).  
GN TNFRSF11B OR OPG.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=INTESTINE;  
RX MEDLINE=97262071; PubMed=9108485;  
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,  
RA Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,  
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,  
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,  
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,  
RA Suggs S., Boyle W.J.;  
RT \*Osteoprotegerin: a novel secreted protein involved in the regulation  
RT of bone density.\*;  
RL Cell 89:309-319(1997).  
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY  
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY  
CC SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
DR EMBL; U94330; AAB53707.1; -.  
DR INTERPRO; IPR001368; -.  
DR PFAM; PF00020; TNFR\_C6; 4.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
DR PROSITE; PS50050; TNFR\_NGFR\_2; 2.  
DR PRODOM; PD000771; -. 1.  
KW Glycoprotein; Repeat; Cytokine; Signal.  
FT SIGNAL 1 21 BY SIMILARITY.

217	CKPCICTVEV	---SVFFPPKPKDVLITLTPKVTCTVVDISKDDPEVQSFVDDVE	27
244	VHNAKTPREQYNSTRYVSVLTVLHQDLWNGKEYCKVSNKALPAIEKITSKAKGP	303	
274	VHTAQTPREQENSTFRSVELPIMHQDLWNGKEFKCRVNSAAFPAIEKITSKTKGRP	333	
304	REFQVNTLPSRDELTKNOVSLTCLVKGYPSDIAVEWESNGOPENNYKTTTPVLDSDGS	363	
334	KAPQVYTIPTPPKQMAKDKVSTCMITDFPEDITVEWQNGQPAENYKNTQPIMTDGS	393	
364	FFLYSKLITVDKSWQOQNVFSCVSMHEALHNHYTKQSLSLSPG	406	
394	YFYYSKLVQKSNWEAGNFTCSVLHGLHNHHTKANLSHSPG	436	
RESULT	6		
O95407			
ID	O95407	PRELIMINARY;	PRT; 300 AA.
IC	O95407;		
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)	
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)	
DE	DECY RECEPTOR 3 (M68)	(M68C) (M68E).	
GN	DCR3 OR TR6.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates;	Catarrhini; Hominidae; Homo.	
ON	NCBI_TaxID=9606;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=99087326; PubMed=9872321;		
RA	Pidd R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,		
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,		
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,		
RA	Goddard A.D., Botstein D., Ashkenazi A.;		
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and		
RT	colon cancer.";		
RL	Nature 396:699-703(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RP	TISSUE=BLOOD;		
RX	MEDLINE=99253915; PubMed=10318773;		
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;		
RT	"A newly identified member of tumor necrosis factor receptor		
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis.";		
RN	J. Biol. Chem. 274:13733-13736(1999).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RP	TISSUE=PANCREAS;		
RX	MEDLINE=20122600; PubMed=10655513;		
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,		
RA	Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;		
RT	"Overexpression of M68/DCR3 in human gastrointestinal tract tumors		
RT	independent of gene amplification and its location in a four-gene		
RT	cluster.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).		
DR	EMBL; AF104419; AAD03056.1; -		
DR	EMBL; AF134240; AAD29688.1; -		
DR	EMBL; AF217796; AAF35244.1; -		
DR	EMBL; AF217793; AAF33685.1; -		
DR	EMBL; AF217794; AAF33686.1; -		
DR	HSSP; P25942; ICDF.		
DR	INTERPRO; IPR000561; -		
DR	INTERPRO; IPR001368; -		
DR	PFAM; PF00020; TNFR_c6; 4.		
DR	PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.		
DR	PROSITE; PS01186; EGF_2; UNKNOWN_1.		
DR	PROSITE; PS50050; TNFR_NGFR_2; 2.		
DR	PRODOM; PD000771; -; 1.		
KW	Receptor.		
SW	SEQUENCE 300 AA, 32679 MW, F90AEE33718449AF CRC64;		

Query Match 19.1%; Score 433.5; DB 4; Length 300;  
 Best Local Similarity 39.1%; Pred. No. 3.1e-30;  
 Matches 72; Conservative 32; Mismatches 69; Indels 11; Gaps 1;

QY 5 PKYHYDEETSHQLLCKDPCPGYLYKHCHTAKWTVCAPCPDHYHYDSDWHTSDECLYCSP 64  
 DB 34 PTPYWRDAETGERLVAQCPGPGTFVQRCRDSFTTCGCPGPRHYTFQWNYLERCRYCNV 93  
 QY 65 VKELQVVKQECNRTHNRVCECKEGRYLEFCLKHRSCPPGFGVVOAGTPERTNVCCKRC 124  
 DB 94 LCGEREEARACHATHNRACRCRTGFAHAGFLEHASCPPGAGVATAGTPSNTQOCPC 153  
 QY 125 PDGFFNETSSKAPCRKHTNCVFGLLLTOKGNATHDNICSG-----NSESTQK 173  
 DB 154 PPGTFASSSSSQCPQHRNCTALGLALNVPGSSSHDTLCTSGTGPLSTRVPGAECER 213  
 QY 174 CGID 177  
 DB 214 AVID 217

RESULT 7  
 Q9PUS0 PRELIMINARY; PRT; 302 AA.  
 AC Q9PUS0;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE DECOY RECEPTOR.  
 OS Salvelinus fontinalis (Brook trout).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.  
 OX NCBI\_TaxID=8038;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bobe J., Goetz F.W.;  
 RT "A tumor necrosis factor receptor homolog is up-regulated in the brook trout (Salvelinus fontinalis) ovary at the completion of ovulation."  
 RL Biol. Reprod. 0:0-0(1999).  
 DR EMBL; AF156738; AD56428.1; -.  
 DR HSSP; P19438; IEXT.  
 DR INTERPRO; IPR000561; -.  
 DR INTERPRO; IPR001368; -.  
 DR PFAM; PF00020; TNFR\_C6; 4.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; UNKNOWN\_1.  
 DR PROSITE; PS00050; TNFR\_NGFR\_2; 1.  
 KW Receptor.  
 SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match 17.8%; Score 405.5; DB 13; Length 302;  
 Best Local Similarity 45.2%; Pred. No. 8.9e-28;  
 Matches 71; Conservative 30; Mismatches 53; Indels 3; Gaps 2;

QY 11 DEETSHQLLCKDPCPGYLYKHCHTAKWTVCAPCPDHYHYDSDWHTSDECLYCSPCKELQ 70  
 DB 27 DRYSGLSIVCDRCPPGYLRAPCSMRKSDCAECPNGAYTEFWNHHSKLCURCS-MCAENQ 85

QY 71 YVKQECNRTHNRVCECKEGRYL--EIEFCLKHRSCPPGFGVVOAGTPERTNVCCKRCPDGF 128  
 DB 86 VVKQECSPSNCECKEGYFKNKYACIKHKECPGYGANTGTGPHQDTECVQOAGF 145

QY 129 FSNETSAPCRKHTNCVFGLLLTOKGNATHDNICS 165  
 DB 146 YSEVSSAKATCLAQSNCKVGLRVLKGDWHTNLCA 182

RESULT 8  
 Q16042 PRELIMINARY; PRT; 439 AA.  
 ID Q16042  
 AC Q16042;

DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE TUMOR NECROSIS FACTOR RECEPTOR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91370690; PubMed=1966549;  
 RA Dembic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R., Brockhaus M., Lesslauer W.;  
 RT "Two human TNF receptors have similar extracellular, but distinct intracellular, domain sequences."  
 RL Cytokine 2:231-237(1990).  
 DR EMBL; S63368; AAB19824.1; -.  
 DR HSSP; P25942; ICDP.  
 DR INTERPRO; IPR001368; -.  
 DR PFAM; PF00020; TNFR\_C6; 4.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; 2.  
 DR PROSITE; PS00050; TNFR\_NGFR\_2; 3.  
 DR PRODOM; PD000771; -.  
 SQ SEQUENCE 439 AA; 46090 MW; FEBCE329CC67FF6 CRC64;

Query Match 15.1%; Score 343.5; DB 4; Length 439;  
 Best Local Similarity 27.3%; Pred. No. 3.8e-22;  
 Matches 110; Conservative 44; Mismatches 162; Indels 87; Gaps 16;

QY 9 HYDEETSHQLLCKDPCPGYLYKHCHTAKWTVCAPCPDHYHYDSDWHTSDECLYCSPVCK 68  
 DB 23 YIDQ--TAQMCCSKCSPGQAHKVCFTKSTVTDVDSCESTYTLNWNVPCSLCGSCRSS 80

QY 69 LQYVKQECNRTHNRVCECKEGRYLEI-----EFLKHRSCPPGFGVVOAGTPERTNVC 122  
 DB 81 DQVETQACTREQNRICTRCPGYCALSKQEGRCALPKRCRPGFVARPGTETSDVVK 140

QY 123 RCPDGFESNETSSKAPCRKHTNCVFGLLLTOKGNATHDNICSNSESTQK-----GI 176  
 DB 141 PCAPGTFSTSTSDICRPHQICNVVAI-----PGNASMDAVCTSTSPTRSMAPGAHLQ 196

QY 177 DVTVDKTHTCP---PCPAPE---LIGGPSVFLFPKPKDT-----LMISRT----- 217  
 DB 197 PVSTRSQHTOPTPEPSTAPSTSLPMGPS-----PPAEGSTGDFALPVLIVGTALCLL 252

QY 218 ---EVTGVVVDVSHEDP-----EVKFNWYVDGVEVHNK-TKPREQYNTYRVSVLT 268  
 DB 253 IIGVWNCVIMTVKKKPLCLQREAK-----VPHLPADKARGTQGPQOH-----LLIT 300

QY 269 LHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTK----- 320  
 DB 301 APSSSSSLESSASALDRRAP-----TRNQAPAGVEASGAGARASTGSSDSPG 351

QY 321 ---NOVSLTCLVKGFPSPDIAVEWESN-----GOPENNYKTT 355  
 DB 352 GHGTQVNTCIVNVCSHSSDSSSOASSTMGDTDSSPESP 394

RESULT 9  
 Q62327 PRELIMINARY; PRT; 459 AA.  
 ID Q62327  
 AC Q62327;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=NOD;  
RA MEDLINE=95178848; PubMed=7873884;  
RX Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;  
RT "Allelic variation of the type 2 tumor necrosis factor receptor  
gene.";

RL Mamm. Genome 5:726-727(1994).

DR EMBL; X76401; CAA53981.1; -

DR HSP; P19438; INCF.

DR INTERPRO: IPR001368; -

DR PFAM: PF00020; TNFR\_C6; 4

DR PROSITE; PS00652; TNFR\_NGFR\_1; 2

DR PROSITE; PS00050; TNFR\_NGFR\_2; 3

FT NON\_TER 1 1

FT VARIANT 87 87 S -> T.

FT VARIANT 93 93 T -> I.

FT VARIANT 268 268 F -> I.

FT VARIANT 345 345 S -> F.

FT VARIANT 421 421 Y -> C.

SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match 15.08; Score 342; DB 11; Length 459;

Best Local Similarity 25.84; Pred. No. 5.4e-22;

Matches 103; Conservative 52; Mismatches 166; Indels 78; Gaps 13;

Qy 9 HYDETSHOLLCKDCPCPGTYLKHQCTAKWKTVCAPCPDHYVYDTSWHTSDECLYCSPVCKE 68

Db 31 YDRKA--QMCCAKCPGQYVYKFCNKTSDIVCADCEASMTQVWNOFRTCLSCSSCGST 88

Qy 69 LOYVQKQECNRTHNRVCEKGRYLEIEF-----CLKHRSCTPPGFGVVGAGTPTNVC 121

Db 89 DQVETRACTQONRVCAEAGRYCALKTHSGRCQCMRLSKCGPGFGVASSRAPNGVLC 148

Qy 122 KRCPDGPFSTSSKAPCRKHTNCSVFGLLLTQKGNATHDNCNSESSTOKCGIDVTV- 180

Db 149 KACAPGTFSDTSDVCRPHRICSLAI-----PGNASTDAVCA--PESPTLSAIPRTLY 202

Qy 181 ----DKHTCP-----PCPAPELL-----GGPSV-----FLFP 204

Db 203 VSQPEPTRSQPLDQEPGSPQPSILTSGLSTPIEQSTKGISLPIGLIVGTVSLGLML 262

Qy 205 PKPKDTLMISTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVS 264

Db 263 GLVNCFLVQRKKPSCQLQDA--KVHPVDEKSDAVGL-----EQQH----- 304

Qy 265 VLTVLHQLWLNKKEYCKVSKNALPA---PIEKTISKAGQPREPOVYTLPPSRDELTKN 321

Db 305 LLTAPSSSSSSLESSASAGDRAPPGGHPQARVMAEAGSQEASRRISDSHSGHGT 364

Qy 322 QVSLTCLVKGFPSPDIAVEWESN-----GOPENNYKTP 355

Db 365 HNVNCTIVNVCSSDHSQCSQASATVGDPAKPSASP 403

RESULT 10

O88734

ID

AC

DT

DT

DT

DE

GN

OS

OC

OC

OX

RN

RP

RA

RT

RT

RL

DR EMBL; Y14619; CAA74969.1; -

DR EMBL; Y14620; CAA74969.1; JOINED.

DR EMBL; Y14621; CAA74969.1; JOINED.

DR EMBL; Y14622; CAA74969.1; JOINED.

DR EMBL; Y14623; CAA74969.1; JOINED.

DR EMBL; Y14679; CAA74969.1; JOINED.

DR HSP; P19438; INCF.

DR INTERPRO: IPR001368; -

DR PFAM: PF00020; TNFR\_C6; 4

DR PROSITE; PS00652; TNFR\_NGFR\_1; 2

DR PROSITE; PS00050; TNFR\_NGFR\_2; 3

DR PRODOM; PD000771; -; 1

SQ SEQUENCE 482 AA; 51106 MW; F6C15046B48FF83C CRC64;

Query Match 14.78; Score 335; DB 11; Length 482;

Best Local Similarity 25.78; Pred. No. 2.3e-21;

Matches 105; Conservative 49; Mismatches 166; Indels 88; Gaps 14;

Qy 9 HYDETSHOLLCKDCPCPGTYLKHQCT-----AKWTVCAPCPDHYVYDTSWHTSDECLY 61

Db 46 YDRKA--QMCCAKCPGQYVYKFCNKTSDIVCADCEASMTQVWNOFRTCL 103

Qy 62 CSPVCKELQYVQKQECNRTHNRVCEKGRYLEIEF-----CLKHRSCTPPGFGVVGAGT 114

Db 104 CSSSCSTDDQVETRACTQONRVCAEAGRYCALKTHSGRCQCMRLSKCGPGFGVASSRA 163

Qy 115 PERNTVCKRCPDGFSTSSKAPCRKHTNCSVFGLLLTQKGNATHDNCNSESSTOKC 174

Db 164 PNGVLCACAPGTFSDTSDVCRPHRICSLAI-----PGNASTDAVCA--PESPTLS 217

Qy 175 GIDVTV-----DKHTCP-----PCPAPELL-----GGPSV----- 200

Db 218 AIPRTLYVSQPEPTRSQPLDQEPGSPQPSILTSGLSTPIEQSTKGISLPIGLIVGT 277

Qy 201 ---FLFPKPKDTLMISTPEVTCVVDV--VSHEDPEVKFNWYDGVGVHNAKTKPREEQ 255

Db 278 SLGLMLGLVNCFLVQRKKPSCQLQDAKVPVH-PDEKSDAVGLEQHLHTTAPSSSS 336

Qy 256 YNSYRVVSVLTVLHQLWLNKKEYCKVSKNALPA---PIEKTISKAGQPREPOVYTL 312

Db 337 SSSL-----ESSASAGDRAPPGGHPQARVMAEAGSQEASRRIS 378

Qy 313 PSRDELTKNOVSLTCLVKGFPSPDIAVEWESN-----GOPENNYKTP 355

Db 379 DSSHGSHGTHVNTCTIVNVCSSDHSQCSQASATVGDPAKPSASP 426

RESULT 11

O75509

ID

AC

DT

DT

DT

DE

DE

GN

OS

OC

OC

OX

RN

RP

RA

RA

RT

RT

RL

RL

RL

RL

RL

RL

RL

DR EMBL; X76401; CAA53981.1; -

DR HSP; P19438; INCF.

DR INTERPRO: IPR001368; -

DR PFAM: PF00020; TNFR\_C6; 4

DR PROSITE; PS00652; TNFR\_NGFR\_1; 2

DR PROSITE; PS00050; TNFR\_NGFR\_2; 3

FT NON\_TER 1 1

FT VARIANT 87 87 S -> T.

FT VARIANT 93 93 T -> I.

FT VARIANT 268 268 F -> I.

FT VARIANT 345 345 S -> F.

FT VARIANT 421 421 Y -> C.

SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match 15.08; Score 342; DB 11; Length 459;

Best Local Similarity 25.84; Pred. No. 5.4e-22;

Matches 103; Conservative 52; Mismatches 166; Indels 78; Gaps 13;

Qy 9 HYDETSHOLLCKDCPCPGTYLKHQCTAKWKTVCAPCPDHYVYDTSWHTSDECLYCSPVCKE 68

Db 31 YDRKA--QMCCAKCPGQYVYKFCNKTSDIVCADCEASMTQVWNOFRTCLSCSSCGST 88

Qy 69 LOYVQKQECNRTHNRVCEKGRYLEIEF-----CLKHRSCTPPGFGVVGAGTPTNVC 121

Db 89 DQVETRACTQONRVCAEAGRYCALKTHSGRCQCMRLSKCGPGFGVASSRAPNGVLC 148

Qy 122 KRCPDGPFSTSSKAPCRKHTNCSVFGLLLTQKGNATHDNCNSESSTOKCGIDVTV- 180

Db 149 KACAPGTFSDTSDVCRPHRICSLAI-----PGNASTDAVCA--PESPTLSAIPRTLY 202

Qy 181 ----DKHTCP-----PCPAPELL-----GGPSV-----FLFP 204

Db 203 VSQPEPTRSQPLDQEPGSPQPSILTSGLSTPIEQSTKGISLPIGLIVGTVSLGLML 262

Qy 205 PKPKDTLMISTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVS 264

Db 263 GLVNCFLVQRKKPSCQLQDA--KVHPVDEKSDAVGL-----EQQH----- 304

Qy 265 VLTVLHQLWLNKKEYCKVSKNALPA---PIEKTISKAGQPREPOVYTLPPSRDELTKN 321

Db 305 LLTAPSSSSSSLESSASAGDRAPPGGHPQARVMAEAGSQEASRRISDSHSGHGT 364

Qy 322 QVSLTCLVKGFPSPDIAVEWESN-----GOPENNYKTP 355

Db 365 HNVNCTIVNVCSSDHSQCSQASATVGDPAKPSASP 403

RESULT 10

O88734

ID

AC

DT

DT

DT

DE

GN

OS

OC

OC

OX

RN

RP

RA

RT

RT

RL

RL

RL

DR EMBL; X76401; CAA53981.1; -

DR HSP; P19438; INCF.

DR INTERPRO: IPR001368; -

DR PFAM: PF00020; TNFR\_C6; 4

DR PROSITE; PS00652; TNFR\_NGFR\_1; 2

DR PROSITE; PS00050; TNFR\_NGFR\_2; 3

FT NON\_TER 1 1

FT VARIANT 87 87 S -> T.

FT VARIANT 93 93 T -> I.

FT VARIANT 268 268 F -> I.

FT VARIANT 345 345 S -> F.

FT VARIANT 421 421 Y -> C.

SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match 15.08; Score 342; DB 11; Length 459;

Best Local Similarity 25.84; Pred. No. 5.4e-22;

Matches 103; Conservative 52; Mismatches 166; Indels 78; Gaps 13;

Qy 9 HYDETSHOLLCKDCPCPGTYLKHQCTAKWKTVCAPCPDHYVYDTSWHTSDECLYCSPVCKE 68

Db 31 YDRKA--QMCCAKCPGQYVYKFCNKTSDIVCADCEASMTQVWNOFRTCLSCSSCGST 88

Qy 69 LOYVQKQECNRTHNRVCEKGRYLEIEF-----CLKHRSCTPPGFGVVGAGTPTNVC 121

Db 89 DQVETRACTQONRVCAEAGRYCALKTHSGRCQCMRLSKCGPGFGVASSRAPNGVLC 148

Qy 122 KRCPDGPFSTSSKAPCRKHTNCSVFGLLLTQKGNATHDNCNSESSTOKCGIDVTV- 180

Db 149 KACAPGTFSDTSDVCRPHRICSLAI-----PGNASTDAVCA--PESPTLSAIPRTLY 202

Qy 181 ----DKHTCP-----PCPAPELL-----GGPSV-----FLFP 204

Db 203 VSQPEPTRSQPLDQEPGSPQPSILTSGLSTPIEQSTKGISLPIGLIVGTVSLGLML 262

Qy 205 PKPKDTLMISTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVS 264

Db 263 GLVNCFLVQRKKPSCQLQDA--KVHPVDEKSDAVGL-----EQQH----- 304

Qy 265 VLTVLHQLWLNKKEYCKVSKNALPA---PIEKTISKAGQPREPOVYTLPPSRDELTKN 321

Db 305 LLTAPSSSSSSLESSASAGDRAPPGGHPQARVMAEAGSQEASRRISDSHSGHGT 364

Qy 322 QVSLTCLVKGFPSPDIAVEWESN-----GOPENNYKTP 355

Db 365 HNVNCTIVNVCSSDHSQCSQASATVGDPAKPSASP 403

RESULT 10

O88734

ID

AC

DT

DT

DT

DE

GN

OS

OC

OC

OX

RN

RP

RA

RT

RT

RL

RL

RL

DR EMBL; X76401; CAA53981.1; -

DR HSP; P19438; INCF.

DR INTERPRO: IPR001368; -

DR PFAM: PF00020; TNFR\_C6; 4

DR PROSITE; PS00652; TNFR\_NGFR\_1; 2

DR PROSITE; PS00050; TNFR\_NGFR\_2; 3

FT NON\_TER 1 1

FT VARIANT 87 87 S -> T.

FT VARIANT 93 93 T -> I.

FT VARIANT 268 268 F -> I.

FT VARIANT 345 345 S -> F.

FT VARIANT 421 421 Y -> C.

SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match 15.08; Score 342; DB 11; Length 459;

Best Local Similarity 25.84; Pred. No. 5.4e-22;

Matches 103; Conservative 52; Mismatches 166; Indels 78; Gaps 13;

Qy 9 HYDETSHOLLCKDCPCPGTYLKHQCTAKWKTVCAPCPDHYVYDTSWHTSDECLYCSPVCKE 68

Db 31 YDRKA--QMCCAKCPGQYVYKFCNKTSDIVCADCEASMTQVWNOFRTCLSCSSCGST 88

Qy 69 LOYVQKQECNRTHNRVCEKGRYLEIEF-----CLKHRSCTPPGFGVVGAGTPTNVC 121

Db 89 DQVETRACTQONRVCAEAGRYCALKTHSGRCQCMRLSKCGPGFGVASSRAPNGVLC 148

Qy 122 KRCPDGPFSTSSKAPCRKHTNCSVFGLLLTQKGNATHDNCNSESSTOKCGIDVTV- 180

Db 149 KACAPGTFSDTSDVCRPHRICSLAI-----PGNASTDAVCA--PESPTLSAIPRTLY 202

Qy 181 ----DKHTCP-----PCPAPELL-----GGPSV-----FLFP 204

Db 203 VSQPEPTRSQPLDQEPGSPQPSILTSGLSTPIEQSTKGISLPIGLIVGTVSLGLML 262

Qy 205 PKPKDTLMISTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVS 264

Db 263 GLVNCFLVQRKKPSCQLQDA--KVHPVDEKSDAVGL-----EQQH----- 304

Qy 265 VLTVLHQLWLNKKEYCKVSKNALPA---PIEKTISKAGQPREPOVYTLPPSRDELTKN 321

Db 305 LLTAPSSSSSSLESSASAGDRAPPGGHPQARVMAEAGSQEASRRISDSHSGHGT 364

Qy 322 QVSLTCLVKGFPSPDIAVEWESN-----GOPENNYKTP 355

</

```
DR EMBL; AF068868; AAC34583.1; -
DR EMBL; AL096801; CAB75692.1; -
DR HSSP; P07174; INGR.
DR INTERPRO; IPR000488; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_6; 4.
DR PFAM; PF00331; death; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00117; DEATH_DOMAIN; 1.
DR PROSITE; PS00050; TNFR_NGFR_2; 1.
DR PRODOM; PD000771; -; 1.
KW Receptor.
SQ SEQUENCE 655 AA; 71844 MW; 48939391C4852A33 CRC64;

Query Match 12.9%; Score 294; DB 4; Length 655;
Best Local Similarity 27.8%; Pred. No. 1.3e-17;
Matches 87; Conservative 44; Mismatches 134; Indels 48; Gaps 11;

QY 7 YLHYDEETSHOLLCDCPCPGTYLKHCTAKWTKVACPCPDHYHYTDSWHTSDECLYCSVPC 66
Db 54 YRHVDRTATGOVLTCDKCPAGTYSEHCTNTSLRVCSGPCVGTFRHENGIEKCHDCSQPC 113
QY 67 KELQYVKEQNRHNRVCECKEGRYLEIEFCLKHRSCTPGFGVVGQAGTPERNVCKRCPD 126
Db 114 PWNIEKLPALCAALTDRECTCPGFMFSQSNATCAPHTVCPVGWGVKRGKGTETEDVRCQCAR 173
QY 127 GFPSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNICS---CNSBSTQKCGIDVTVDKT 193
Db 174 GTFSDVPSSVMKCAKVTDCLSQNLVWIKPGTKETDNVCGTLPFSFSTS----- 222
QY 184 HTCPCPAPPELLGGSPVFLPPPKDGLMISRTPEVTCVVVDVSHED-----PEVKF 235
Db 223 -----PSP-----GTAIF---PRP-EMHETHEVPSSTYVPKGMNSESNSASVRKV-L 267
QY 236 NWTVDGVEVHNATKPREEQNSTYRVSVTLVTHODWLNKGYCKVCSNKAIPAPI--- 292
Db 268 SSTQEGTVPDNTSSARGKEDVNKT--LPNLQVYVNHQ---QGPHRHIL--KLLPSMEATG 320
QY 293 -EKTISKAKQQR 304
Db 321 GEKSTPIKPKR 333

RESULT 12
Q9UP60 PRELIMINARY; PRT; 384 AA.
AC Q9UP60;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SNC73 PROTEIN.
GN SNC73.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng S., Cao J., Cao W., Cai X., Geng L.;
RT "Identification and characterization of SNC73, a gene which is down-
RT regulated in colorectal cancer.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067420; AAC19365.1; -
DR HSSP; P01825; 7FAB.
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; ig; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;

Query Match 12.2%; Score 277.5; DB 4; Length 384;
Best Local Similarity 28.2%; Pred. No. 2e-16;
```

```
Matches 96; Conservative 35; Mismatches 138; Indels 71; Gaps 14;

QY 115 PERNTVCKRCPCDGFESNE-----TSSKAPCRKHTNCSVFGLLLTOK 155
Db 49 PDGNVVIACLVQGFQFFQEPISVTSWESGGQGVTAARNFPSPQDASGLYTTSSQLTLPATQ- 107
QY 156 GNATHDNICSGNSESTOKCGI-----DVTVDKHTCTCPAPPELLGGSPVFLFPKP 207
Db 108 -----CLAGKSVT--CHVKHYTNPSQDVT-----PCVPSPPTPSPST-PPTP 149
QY 208 K-----DTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPKR 252
Db 150 SPSCCHPRLSLHRPALEDLLLGSEANLTCLTGL-RDASGVTTWTPTSSGK--SAVQGP 206
QY 253 EEOYNSTYRVSVTLVTHODWLNKGYCKVCSNKAIPAPIEKTISKAKGQPREPOVYTL 312
Db 207 ERLCGCYSVSSVLPGCAEPWNHGKTFCTAAYPESTKPLTATLSKS-GNTFRFEVHLLP 265
QY 313 PSRDELTKNQ-VSLTCLVKGFPYSDIAVEWESNQ--PENNYKTTTPVLD--SDG--SFFL 366
Db 266 PPEELALNELVLTCLARGFSPKDLVLRWLQSQELPREKYLTWASRQSPSQGTTTFAV 325
QY 367 YSKLTVDKSRMQGNVFCSCVMHEALHNHYTKQSLSLSPG 406
Db 326 TSILRAAEDWKKGDTFCMVHGHEALPLAFTQKTIDRLAG 365

RESULT 13
Q90544 PRELIMINARY; PRT; 684 AA.
AC Q90544;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBI_TaxID=7801;
RN [1]
RP SEQUENCE FROM N.A.
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,
RA Flajnik M.F.;
RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks.";
RL Nature 374:168-173(1995).
DR EMBL; U18701; AAB48195.1; -
DR HSSP; P01857; IFC1.
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; ig; 6.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Signal.
FT SIGNAL.
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;

Query Match 12.1%; Score 275.5; DB 13; Length 684;
Best Local Similarity 26.3%; Pred. No. 5.8e-16;
Matches 101; Conservative 53; Mismatches 161; Indels 69; Gaps 17;

QY 51 DSWHTSDECLYCS-----PVCKELQYVKEQNRHNRVCECKEGRYLEIEFCLKHRS 102
Db 315 EEWQSGVE-YTCSAKQDQSSSTPVVKTRKARVEPTKPHRL----- 354
QY 103 CPFGFGVVGQAGTPERNVCKRCPDGPFNSNETS---SKAPCRKHTNCSVFGLLLTOKNAT 159
Db 355 LPPSPBIEQTSSTATLTCLIR---GFYPDKVSVSWQDDVSVSANVTNFTALEQ--DLT 409
QY 160 HDNICSNSESTQ-KCGIDVTVDKTHTCPP-----CPAPELGGSPSVFLF 203
```



Db 48 CSCEPGKYLSSKCTPTSDVCLPCGPDYLDTWNEEDKCL-LHKVCDAGKALVAV-DPG 105  
QY 77 NRTHNRVCECKEGRY--LEIEFCLKHRCPPGPGVVQAGTPERTVCKRCPDGFFSNETS 134  
Db 106 NHTAPRCACACTAGYHNSDCECCRRNTECAPGEGAGHPLQLNKDTVCTPCLLGGFFSDVFS 165  
QY 135 SKAPCRKHTNCVFGLLLTQKNATHDNTICSGNSESTQKCGIDVVDKTHTCPPCPAPEL 194  
Db 166 STDKCKPWTNCTLLGKLEAHQGTESDVVCS--MTLRRPPKEAQAY 211  
QY 195 LGGPSVLEFPKPKDPLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREE 254  
Db 212 L--PSLIVL-----LLFISVVVAAIIFGV-----YVRKGGK----- 241  
QY 255 QYNSTYRVVSVLTVLHODWLN-----GKEY---KCKVSNKA-----LPAP 291  
Db 242 -----ALTANLWNVNDACSSLSGNKSSGDRGAGSHSATSSQEVCEGILLMTR 291  
QY 292 IEKTISK-----AKGQP-----REPQVYTL-----PSRDELT--KNO 322  
Db 292 EEKWPEDGAGVGPVCAAGGPAEVRDSRTFTLVSEVETQGLSRKIPTEDDYTDORPSQ 351  
QY 323 VSLTCLV-----KGFYPSDIAVEWESNGOPENNYKTTTPVLDSGSPF 365  
Db 352 PSTGSLLLIQGSKSIPFPQEPLEVGENDSLSQCFGTGTSTVDSEGDF 400

Search completed: March 1, 2001, 09:20:09  
Job time: 407 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:45 ; Search time 40.97 Seconds  
(without alignments)  
320.024 Million cell updates/sec

Title: US-09-389-782A-6  
Perfect score: 2274  
Sequence: 1 ETFFPKYLHYDETSKQLC.....VMHEALHNHYTKSLSPG 406

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	1231	54.1	330	1	GCL_HUMAN
2	1148	50.5	326	1	GC2_HUMAN
3	1138.5	50.1	327	1	GC4_HUMAN
4	1134.5	49.9	290	1	GC3_HUMAN
5	917	40.3	323	1	GC_RABIT
6	895.5	39.4	329	1	GC2_CAVPO
7	848.5	37.3	329	1	GC3_MOUSE
8	842.5	37.0	398	1	GC3N_MOUSE
9	834	36.7	333	1	GC3_RAT
10	817.5	35.9	326	1	GCL_MOUSE
11	812.5	35.7	324	1	GCL_RAT
12	812.5	35.7	393	1	GC1N_MOUSE
13	805	35.4	329	1	GCC_RAT
14	804	35.4	330	1	GCA_MOUSE
15	804	35.4	399	1	GCAM_MOUSE
16	797	35.0	335	1	GCAB_MOUSE
17	782	34.4	322	1	GCA_RAT
18	779.5	34.3	326	1	GCB_MOUSE
19	779.5	34.3	405	1	GCBN_MOUSE
20	366.5	16.1	421	1	EPC_MOUSE
21	361.5	15.9	391	1	MUCB_HUMAN
22	359	15.8	454	1	MUC_HUMAN
23	355.5	15.6	429	1	EPC_RAT
24	354	15.6	455	1	MUC_MOUSE
25	332	15.5	428	1	EPC_HUMAN
26	349	15.3	476	1	MUCN_MOUSE
27	346.5	15.2	458	1	MUC_RABIT
28	345.5	15.2	461	1	TNR2_HUMAN
29	341.5	15.0	479	1	MUCM_RABIT
30	337	14.8	474	1	TNR2_MOUSE
31	335.5	14.8	450	1	MUC_CANFA
32	330.5	14.5	454	1	MUC_MESAU
33	329.5	14.5	457	1	MUC_SUNNU

34	308.5	13.6	438	1	HVC2_HETFR	P23085 heterodontu
35	296	13.0	438	1	HVC5_HETFR	P23087 heterodontu
36	295	13.0	299	1	ALC_RABIT	P01879 oryctolagus
37	292.5	12.9	446	1	MUC_CHICK	P01875 gallus gall
38	292	12.8	461	1	HVCM_HETFR	P23088 heterodontu
39	291.5	12.8	289	1	CD40_MOUSE	P27512 mus musculu
40	285	12.5	393	1	HVC3_HETFR	P23086 heterodontu
41	277.5	12.2	353	1	ALC1_HUMAN	P01876 homo sapien
42	276.5	12.2	353	1	ALC1_GORGO	P20758 gorilla gor
43	276.5	12.2	370	1	HVC1_HETFR	P23084 heterodontu
44	273.5	12.0	340	1	ALC2_HUMAN	P01877 homo sapien
45	264	11.6	277	1	CD40_HUMAN	P25942 homo sapien

## ALIGNMENTS

RESULT	1
GCL_HUMAN	
ID	GCL_HUMAN
AC	P01857
DT	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	IG GAMMA-1 CHAIN C REGION.
GN	IGHG1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=82274238; PubMed=6287432;
RA	Ellison J.W., Berson B.J., Hood L.E.;
RT	"The nucleotide sequence of a human immunoglobulin C gamma gene."
RL	Nucleic Acids Res. 10:4071-4079(1982).
RN	[2]
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
RX	MEDLINE=71064024; PubMed=5489771;
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
RA	Waxdal M.J., Edelman G.M.;
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino
RT	acid sequence of heavy-chain cyanogen bromide fragments H1-H4."
RL	Biochemistry 9:3161-3170(1970).
RN	[3]
RP	SEQUENCE OF 136-329 (EU).
RX	MEDLINE=71064025; PubMed=5530842;
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.;
RA	Edelman G.M.;
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino
RT	acid sequence of heavy-chain cyanogen bromide fragments H5-H7."
RL	Biochemistry 9:3171-3181(1970).
RN	[4]
RP	SEQUENCE (MYELOMA PROTEIN NIE).
RX	MEDLINE=77070269; PubMed=826475;
RA	Ponstingl H., Hilschmann N.;
RT	"The rule of antibody structure. The primary structure of a
RT	monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The
RT	chymotryptic peptides of the H-chain, alignment of the tryptic
RT	peptides and discussion of the complete structure."
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN	[5]
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
RX	MEDLINE=83289131; PubMed=6884994;
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT	"Three-dimensional structure determination of antibodies. Primary
RT	structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN	[6]
RP	DISULFIDE BONDS.
RX	MEDLINE=71064027; PubMed=4923144;
RA	Gall W.E., Edelman G.M.;
RT	"The covalent structure of a human gamma G-immunoglobulin. X.
RT	Intrachain disulfide bonds."





```
DR INTERPRO: IPR003006; -
DR PFAM: PF00047; Ig; 3.
DR PROSITE: PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 220 CH2.
FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;

Query Match 50.1%; Score 1138.5; DB 1; Length 327;
Best Local Similarity 71.1%; Pred. No. 4.0e-70;
Matches 224; Conservative 16; Mismatches 30; Indels 45; Gaps 5;

QY 137 APCRKHN-----CSV-----FGLLLTKGNATHDNCISG 166
DB 12 APCRSYSESTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLOSLSSVTV 71
QY 167 NSES-----TQCGID-----VTVDKT-----HTCPCPAPELGGPSVFLFPKPKDTL 211
DB 72 PSSSLGTYTCNVYDHKPSNTKVDKRVESKYGPPCPAPPEFVGGSVFLFPKPKDTL 131
QY 212 MISRTPEVTCVVVDVSHEDPEVKFNWYGVGVEVHNAKTKPREQYNSTYRVVSVLTVHLQ 271
DB 132 MISRTPEVTCVVVDVSEDEPEVQFNWYGVGVEVHNAKTKPREQYNSTYRVVSVLTVHLQ 191
QY 272 DWLNGKEYKCKVSKNALKPAIEKTIISKAKGPQEPQVYTLPPSDELTKNOVSTCLVKG 331
DB 192 DWLNGKEYKCKVSKNGLPSSIEKTIISKAKGPQEPQVYTLPPSDEMTKNOVSTCLVKG 251
QY 332 FYPSDIAVEWESNGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQGNVFSCSVHMEA 391
DB 252 FYPSDIAVEWESNGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQGNVFSCSVHMEA 311
QY 392 LHNHYTKSLSPG 406
DB 312 LHNHYTKSLSLG 326

RESULT 4
GC3_HUMAN STANDARD; PRT; 290 AA.
AC P01860;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).
GN IGHG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RN FRAGMENT (DISEASE PROTEIN WIS).
RX MEDLINE-81021548; PubMed-6774747;
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
RL Biochemistry 19:4304-4308(1980).
[2]
RN REVISIONS TO 12-97 OF PROTEIN WIS.
RX MEDLINE-77118561; PubMed-402363;
RA Michaelson T.E., Frangione B., Franklin E.C.;
RT "Primary structure of the 'hinge' region of human IgG3. Probable
RL quadruplication of a 15-amino acid residue basic unit.";
Biochem. Biophys. Res. Commun. 71:907-914(1976).
[3]
RN REVISIONS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).
RX MEDLINE-77021516; PubMed-823945;
RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
RL Structure of the FC fragment of immunoglobulin G3.";
Biochem. Biophys. Res. Commun. 71:907-914(1976).
[4]
RN SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
RX MEDLINE-82247835; PubMed-6808505;
RA Alexander A., Steinmetz M., Barritault D., Frangione B.,
RL Franklin E.C., Hood L., Buxbaum J.N.;
RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
RL gene deletion model.";
Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
CC NORMALLY PRESENT IN THE HINGE REGION.
CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
CC REF. 2.
CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
CC AND ALL OF THE CH1 REGION.
CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
CC GAMMA-3 HEAVY CHAINS.
CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
CC OR ANOTHER GAMMA CHAIN SUBCLASS.
CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
CC SEGMENT (12-28).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: J00231; AAA52805.1; ALT_SEQ.
PIR: A02149; G3HUWI.
MIM: 147120;
DR INTERPRO: IPR000495; -
DR INTERPRO: IPR003006; -
DR PFAM: PF00047; Ig; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT DOMAIN 12 73 HINGE.
FT DOMAIN 74 183 CH2.
FT DOMAIN 184 289 CH3.
FT REPEAT 29 43
FT REPEAT 44 58
FT REPEAT 59 73
FT MOD_RES 1 1
FT CARBOHYD 6 6
FT DISULFID 7 7 N-LINKED (GLCNAC. . .).
FT DISULFID 24 24 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 27 27 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 33 33 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 39 39 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 42 42 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 48 48 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 54 54 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 57 57 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 63 63 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 69 69 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT DISULFID 72 72 INTERCHAIN (WITH HEAVY CHAIN DIMER).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .).
FT MOD_RES 290 290 REMOVED POST-TRANSLATIONALLY.
FT VARIANT 126 127 QV -> EB (IN ZUC).
FT FTID-VAR_003890.
```







```
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 36.7%; Score 834; DB 1; Length 333;
Best Local Similarity 55.6%; Pred. No. 1.9e-49;
Matches 160; Conservative 41; Mismatches 55; Indels 32; Gaps 5;

QY 130 SNETSSKAPCRKTKNCSVFGLLLTQKNATHDNICNSSESTQCGIDVAVDKTH-----184
DB 66 SSVTSSTWPSQTWT-C-----NVAHP-----ASTK---VDKKVERNGGIGH 104

QY 185 -----TCPPCAPPELLGGPSVFLFPKPKDLMISRTPEVTCVVDVSHEDPEVKFNWY 238
DB 105 KCPTCTCHKCPVPELLGGPSVFLFPKPKDILLISQNAKVTVCVVVDVSEEDPDVQSF 164

QY 239 VDGVEVINAKTKPREQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISK 298
DB 165 VNNVEVHTAQTPREQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPSPIEKTISK 224

QY 299 AKGQPREQYVTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPPENNYKTTTPVL 358
DB 225 PKGLVRAPQVYVGGPPEQTEQTSVLTCLTSGLPNDIGVETSNGHIEKNKNTPEVM 284

QY 359 DSDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKLSLSLSPG 406
DB 285 DSDGSFFMYSKLVNRSRWDSPRAFCVSVVHGLNHHVKSISRPPG 332

RESULT 10
GC1_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE IG GAMMA-1 CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0017; PS0017.
DR INTERPRO; IPR000495; -.
DR PFAM; PF00047; ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 35.9%; Score 817.5; DB 1; Length 326;
Best Local Similarity 60.5%; Pred. No. 1.9e-48;
Matches 144; Conservative 42; Mismatches 37; Indels 15; Gaps 3;

QY 172 QKCGIDVTVDKTHTCPCPAPELLGG---PSVFLFPKPKDLMISRTPEVTCVVDVSH 228
DB 100 RNCGGD-----CKPC-----ICTGSEVSVFIFPPKPKDVLITLTTPKVTGVVDISQ 147
QY 229 EDEPVKNWYVGVGHNAKTKPREQYNSTYRVSVLTVLHODWLNKGYCKVSNKAL 288
DB 148 DDPEVHFSWFDVDEHTAQRPEEQNFSTFRSVSELPILHODWLNKGRFRCKVTSAAF 207
QY 289 PAPIEKTISKAKGQPREQYVTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPE 348
DB 208 PSPIEKTISKPEGRTOVPHVYVTSPTKEMTQNEVTSITCVKGYFPDPIYVEMQNGQPQ 267
QY 349 NNYKTPVPLDSDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKLSLSLSPG 406
DB 268 ENYKNTPTMTDTCGDFLYSKLVNKKWQGNFTTCSVLHGLNHHHTKSLSHSPG 325

RESULT 11
GC1_MOUSE STANDARD; PRT; 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RX MEDLINE=80202559; PubMed=6769752;
RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
cloned in a bacterial plasmid.";
RL Gene 9:87-97(1980).
RN [3]
RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=80012837; PubMed=113776;
RA Rogers J., Clarke P., Salsner W.;
RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
heavy chain.";
RL Nucleic Acids Res. 6:3305-3321(1979).
RN [4]
RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=78242288; PubMed=98524;
RA Adetugbo K.;
RT "Evolution of immunoglobulin subclasses. Primary structure of a
murine myeloma gamma 1 chain.";
RL J. Biol. Chem. 253:6068-6075(1978).
RN [5]
RP DISULFIDE BONDS (MOPC 21).
RX MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RL Biochem. J. 126:837-850(1972).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; V00793; CAA24172.1; -.
```



```
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR EMBL; V00793; CAA24175.1; -
DR EMBL; V00795; CAA24176.1; -
DR PIR; A02159; GMS.
DR MGD; MGI:96446; IGH-4.
DR INTERPRO; IPR000495; -
DR PFAM; PF00047; I9; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 110
FT DOMAIN 111 217
FT DOMAIN 218 324
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 104 104
FT DISULFID 107 107
FT DISULFID 109 109
FT DISULFID 138 198
FT CARBOHYD 174 174
FT DISULFID 244 302
FT MOD_RES 324 324
FT CONFLICT 276 276
FT CONFLICT 278 278
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 35.7%; Score 812.5; DB 1; Length 324;
Best Local Similarity 62.3%; Pred. No. 4.2e-48;
Matches 139; Conservative 44; Mismatches 35; Indels 5; Gaps 2;

QY 186 CPP--CPAPELGGPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNMYDQVE 243
DB 104 CKPCICTVPEV---SSVFIPPKPKDVLITLTPKVTCCVVVDISKDPEVQFSWFVDVDE 160

QY 244 VHNAKTPREQNSYRVSVLTVLHQLWLNKYEKCKVSNKALPAPIETKISKAKGP 303
DB 161 VHTAQTPREEQNSFRSSELPIMHQLWLNKYEKCRVNSAFAPIETKISKTKGRP 220

QY 304 REPQVTLPPSRDELKNOVSLCLVKGFYPSDIAVWESNGOPENNYKTPPVLDSDGS 363
DB 221 KAPQVTTPPKQMAKDKVSLTCMTDFPEPDTVWQNGQPAENYKNTQPLMNTGS 280

QY 364 FFYLSKLTVDKSRWQGNVFCSSVMHEALNHYTKRSLSLSPG 406
DB 281 YFVYSKLVNOKSNWEAGNTFTCSVLHGLNHHTEKSLSHSPG 323

RESULT 12
GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-80045036; PubMed-115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568 (1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
```

```
RX MEDLINE-82197626; PubMed-6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE-82115295; PubMed-6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE-82222190; PubMed-6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
SEGMENT OF MU CHAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR PIR; B02159; GMSM.
DR MGD; MGI:96446; IGH-4.
DR INTERPRO; IPR000495; -
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; I9; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 110
FT DOMAIN 111 217
FT DOMAIN 218 324
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 104 104
FT DISULFID 107 107
FT DISULFID 109 109
FT DISULFID 138 198
FT CARBOHYD 174 174
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 35.7%; Score 812.5; DB 1; Length 393;
Best Local Similarity 62.3%; Pred. No. 5.1e-48;
Matches 139; Conservative 44; Mismatches 35; Indels 5; Gaps 2;

QY 186 CPP--CPAPELGGPSVFLPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNMYDQVE 243
DB 104 CKPCICTVPEV---SSVFIPPKPKDVLITLTPKVTCCVVVDISKDPEVQFSWFVDVDE 160
```

QY 244 VHNKTRPREQYNSYTRVSVLVLHODWLNKGYCKYVSNKALPAPIEKTISKAKGP 303  
 Db 161 VHTAQTPREQNFSTERSVSELPIMHQDWLNKGYCKYVSNKALPAPIEKTISKGRP 220  
 QY 304 REPOVYTLPPSRDELTKNOVSLTCLVKGFFPSDITAVEWESNGQPNYKTPPVLDSDGS 363  
 Db 221 KAPOVYTIPTPPEQMAKDKVSLTCMIDFFPEDITVEMQWNGQPAENYKTPQIMNTNGS 280  
 QY 364 FFLYSLKTVKDSRWQOQNVFSCVSNHLYTOKLSLSLSPG 406  
 Db 281 YFVYSLKLVQKSNWEAGNTFTCSVHLEGLNHNHTEKLSLSHSPG 323

RESULT 13  
 GCC\_RAT  
 ID GCC\_RAT STANDARD; PRT; 329 AA.  
 AC P20762;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-2C CHAIN C REGION.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88166903; PubMed=3127222;  
 RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;  
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant  
 RL Eur. J. Immunol. 18:317-319(1988).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X07189; CAA30169.1; -  
 DR PIR: S00847; S00847.  
 DR INTERPRO: IPR000495; -  
 DR INTERPRO: IPR003006; -  
 DR PFAM: PF00047; ig; 3.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region.  
 FT NON\_TER 1  
 FT DOMAIN 1 97 CHI.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 222 CH2.  
 FT DOMAIN 223 329 CH3.  
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 143 203  
 FT DISULFID 249 307  
 SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 35.4%; Score 805; DB 1; Length 329;  
 Best Local Similarity 57.7%; Pred. No. 1.4e-47;  
 Matches 146; Conservative 44; Mismatches 49; Indels 14; Gaps 2;

QY 168 SESTQKQIDVTVDKTHTC-----pp---CPAPELLGGPSVFLPPKPKDILMI 213  
 Db 76 SSQVTCVSAHPATYKSLIKRIEPRPKPRPTDICSDDNLGRPSVFIFPPKPKDILMI 135  
 QY 214 SRTPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNAKTPREQYNSTYRVSVLTVLHQDW 273  
 Db 136 TLTKPVTCVVVDVSEEDPDVQSFVFNVRVFTAQTPQHEQLNGTRFVSTLHQDW 195

QY 274 LNGKEYCKYVSNKALPAPIEKTISKAKGPQVYTLPPSRDELTKNOVSLTCLVKGFF 333  
 Db 196 MSGREFKCKVNNKDLPSPIEKTISKPRKARTPOVYTIPTPPEQMSKNKVSILTCMWTSEY 255  
 QY 334 PSDITAVEWESNGQPNYKTPPVLDSDGSFFLYSLKTVKDSRWQOQNVFSCVSNHLY 393  
 Db 256 PASISVEMERGELEQDYKNTPLVLDSDSEYFLYSLKTVKDSRWQOQNVFSCVSNHLY 315  
 QY 394 NHYTOKLSLSLSPG 406  
 Db 316 NHHTQKNLSRSPG 328

RESULT 14  
 GCAA\_MOUSE  
 ID GCAA\_MOUSE STANDARD; PRT; 330 AA.  
 AC P01863;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE IG GAMMA-2A CHAIN C REGION, A ALLELE.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=81076554; PubMed=6777755;  
 RA Sikorav J.-L., Auffray C., Rougeon F.;  
 RT "Structure of the constant and 3' untranslated regions of the murine  
 RL Balb/c gamma 2a heavy chain messenger RNA.";  
 RL Nucleic Acids Res. 8:3143-3155(1980).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=81198976; PubMed=6262729;  
 RA Yanawaki-Kataoka Y., Miyata T., Honjo T.;  
 RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene  
 RT and evolution of heavy chain genes: further evidence for intervening  
 RT sequence-mediated domain transfer.";  
 RL Nucleic Acids Res. 9:1365-1381(1981).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=8123894; PubMed=6787604;  
 RA Ollio R., Auffray C., Mochamps C., Rougeon F.;  
 RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes  
 RT suggests that exons can be exchanged between genes in a multigenic  
 RT family.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).  
 RN [4]  
 RP MYELOMA PROTEIN MOPC 173.  
 RX MEDLINE=74175517; PubMed=4831970;  
 RA Bourgois A., Fougereau M., Rocca-Serra J.;  
 RT "Determination of the primary structure of a mouse IgG2a  
 RT immunoglobulin: amino-acid sequence of the Fc fragment. Implications  
 RT for the evolution of immunoglobulin structure and function.";  
 RL Eur. J. Biochem. 43:423-435(1974).  
 RN [5]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=73056887; PubMed=4565406;  
 RA de Preval C., Fougereau M.;  
 RT "Determination of the primary structure of a mouse gamma G2a  
 RT immunoglobulin. Identification of the disulfide bridges.";  
 RL Eur. J. Biochem. 30:452-462(1972).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: V00798; CAA24178.1; -  
 DR PIR: A02152; G2MSA.

DR INTERPRO: IPR000495; -  
DR INTERPRO: IPR003006; -  
DR PFAM: PF00047; ig; 3.  
DR PROSITE: PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1  
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 82  
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 144 204  
FT DISULFID 250 308  
FT MOD\_RES 330 330 REMOVED POST-TRANSLATIONALLY.  
SQ SEQUENCE 330 AA; 36389 MM; B84361C5445A6864 CRC64;

Query Match 35.4%; Score 804; DB 1; Length 330;  
Best Local Similarity 65.0%; Pred. No. 1.6e-47;  
Matches 145; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

QY 186 CPP--CPAPELLGGPSVLEFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFNMYVDGVE 243  
DB 107 CPPCKCPAPNLLGGPSVLEFPKPKDVLMSLSPITVTCVVVDVSEDDPDVQISFVNNVE 166  
QY 244 VHNAKTKPREEQNSTYRVSVLTVLHQLDNLNGKEYCKVSKNKPAPLPIETISKAKGP 303  
DB 167 VHTAQQTQTHREDYNSLTVVVSALPIQHDWMSGKEFKCKVNNKDLPAPIERTISKPKGSV 226  
QY 304 REPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGS 363  
DB 227 RAPQVYVLPPEEEMTKQVLTLCMTDFMPEDIIYVETNNGKTELNYKNTPEVLDSDGS 286  
QY 364 FFYLSKLTVDKSRWQGNVFCSCVMHEALNNHYTKQSLSPG 406  
DB 287 YFMYSLRVEKKNWERNYSYSCVVEGLNHHHTTKFSRTPG 329

RESULT 15  
GCAM\_MOUSE STANDARD; PRT; 399 AA.  
AC P01865;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2A CHAIN C REGION, MEMBRANE-BOUND FORM.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-8222190; PubMed-6283537;  
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;  
RT "Nucleotide sequences of gene segments encoding membrane domains of  
RT immunoglobulin gamma chains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).  
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA  
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED  
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-  
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED  
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND  
CC SEGMENT OF MU CHAINS.  
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE  
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF  
CC THE A ALLELE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC EMBL: J00471; AAB59661.1; ALT\_INIT.  
DR PIR: A02154; G2MSAM.  
DR MGD: MGI:96443; IGH-1.  
DR INTERPRO: IPR000495; -  
DR INTERPRO: IPR003006; -  
DR PFAM: PF00047; ig; 3.  
DR PROSITE: PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
Transmembrane; Alternative splicing.  
FT NON\_TER 1  
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 82  
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 144 204  
FT DISULFID 250 308  
FT TRANSMEM 346 363  
FT DOMAIN 364 399  
FT CARBOHYD 180 180 N-LINKED (GLCNAC. -) (POTENTIAL).  
SQ SEQUENCE 399 AA; 44020 MM; 4C38138BFAED3FF0 CRC64;

Query Match 35.4%; Score 804; DB 1; Length 399;  
Best Local Similarity 65.0%; Pred. No. 1.9e-47;  
Matches 145; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

QY 186 CPP--CPAPELLGGPSVLEFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFNMYVDGVE 243  
DB 107 CPPCKCPAPNLLGGPSVLEFPKPKDVLMSLSPITVTCVVVDVSEDDPDVQISFVNNVE 166  
QY 244 VHNAKTKPREEQNSTYRVSVLTVLHQLDNLNGKEYCKVSKNKPAPLPIETISKAKGP 303  
DB 167 VHTAQQTQTHREDYNSLTVVVSALPIQHDWMSGKEFKCKVNNKDLPAPIERTISKPKGSV 226  
QY 304 REPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGS 363  
DB 227 RAPQVYVLPPEEEMTKQVLTLCMTDFMPEDIIYVETNNGKTELNYKNTPEVLDSDGS 286  
QY 364 FFYLSKLTVDKSRWQGNVFCSCVMHEALNNHYTKQSLSPG 406  
DB 287 YFMYSLRVEKKNWERNYSYSCVVEGLNHHHTTKFSRTPG 329

Search completed: March 1, 2001, 09:17:47  
Job time: 265 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:46 ; Search time 79.26 seconds  
(without alignments)  
347.813 Million cell updates/sec

Title: US-09-389-782A-6  
Perfect score: 2274  
Sequence: 1 ETPPKYLHYDEETSHOLLC.....VMHEALHHYQKSLSPG 406

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1231	54.1	330	1	GHU	Ig gamma-1 chain C
2	1228	54.0	255	4	S31866	Ig gamma-1 chain C
3	1227	54.0	374	2	S72664	Ig heavy chain V r
4	1222.5	53.8	374	2	S69339	Ig heavy chain V r
5	1180	51.9	234	2	PT0207	Ig gamma chain C r
6	1156	50.8	377	2	A23511	Ig gamma-3 chain C
7	1154	50.7	377	2	A60764	Ig gamma-3 chain C
8	1148	50.5	326	1	G2HU	Ig gamma-2 chain C
9	1138.5	50.1	327	1	G4HU	Ig gamma-4 chain C
10	1134.5	49.9	289	1	G3HUW1	Ig gamma-4 chain C
11	917	40.3	323	1	GHRB	Ig gamma-3 heavy c
12	907	39.9	328	2	I47160	Ig gamma chain C r
13	907	39.9	328	2	I47159	Ig gamma-2b chain
14	905	39.8	277	2	I47162	Ig gamma 2a chain
15	895.5	39.4	329	1	G2GP	Ig gamma 4 chain c
16	891.5	39.2	328	2	I47158	Ig gamma 1 chain c
17	879	38.7	328	2	I47161	Ig gamma 3 chain c
18	858.5	37.8	470	2	G22080	Ig heavy chain pre
19	848.5	37.3	329	1	G3MSC	Ig gamma-3 chain C
20	842.5	37.0	398	1	G3MSM	Ig gamma-3 chain C
21	841	37.0	308	2	C30554	Ig heavy chain C r
22	841	37.0	472	2	S31459	Ig gamma-1 chain -
23	834	36.7	333	2	PS0018	Ig gamma-2b chain
24	822.5	36.2	444	2	PC4436	monoclonal antibod
25	817.5	35.9	326	2	PS0017	Ig gamma-1 chain C
26	812.5	35.7	324	1	G1MS	Ig gamma-1 chain C
27	812.5	35.7	393	1	G1MSM	Ig gamma-1 chain C
28	805	35.4	329	2	S00847	Ig gamma-2c chain
29	804	35.4	330	1	G2NSA	Ig gamma-2a chain

ALIGNMENTS

RESULT 1

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence revision 18-Aug-1982 #text change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Beison, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433; MUID:82274238

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>

A:Cross-references: EMBL:217370

A:Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) marker

A:Note: Lys-330 is removed after translation

R:Harris, L.J.

submitted to the EMBL Data Library, October 1992

A:Reference number: S33904

A:Accession: S36861

A:Molecule type: DNA

A:Residues: 2-330 <CHAR>

A:Cross-references: EMBL:217370

R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaide, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of

A:Reference number: S33887; MUID:83001943

A:Accession: S33887

A:Molecule type: DNA

A:Residues: 88-113,235-330 <TAK>

A:Cross-references: EMBL:217370

R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman

Biochemistry 9, 3161-3170, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq

A:Reference number: A90563; MUID:71064024

A:Contents: myeloma protein Eu

A:Accession: B90563

A:Molecule type: protein

A:Residues: 1-96,'R',98-135 <CUN>

A:Note: this sequence has the G1m(3) marker, 97-Arg

R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se

A:Reference number: A90564; MUID:71064025

A:Contents: Eu

A:Accession: A90564

A:Molecule type: protein

A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2

A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met

R:Ponstingl, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976

A:Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni



Db	75	GVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKKYSNKALPAPIETISKAK	134
Qy	301	GQPEPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDLDS	360
Db	135	GQPEPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDLDS	194
Qy	361	DGSFFLYSKLTVDKSRWQOGNVFSCSVMHREALHNHYTQKS	400
Db	195	DGSFFLYSKLTVDKSRWQOGNVFSCSVMHREALHNHYTQKS	234
RESULT	6		
Ig gamma-3 chain C region (allotype G3m(b)) - human			
C:Species:	homo sapiens (man)		
C:Date:	28-Dec-1987	#sequence_revision	28-Dec-1987
C:Accession:	A23511	#text_change	23-Jul-1991
Nucleic Acids Res. 14, 1779-1789, 1986			
A:Title:	Sequence of a human immunoglobulin gamma 3 heavy chain constant r		
A:Reference number:	A23511; MUID:86148507		
A:Accession:	A23511		
A:Molecule type:	DNA		
A:Residues:	1-377 <HUC>		
A:Cross-references:	GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID		
C:Genetics:			
A:Gene:	GDB:IGHG3		
A:Cross-references:	GDB:119339; OMIM:147120		
A:Map position:	14q32.33-14q32.33		
A:Introns:	98/3; 115/3; 130/3; 145/3; 160/3; 270/3		
C:Superfamily:	immunoglobulin C region; immunoglobulin homology		
C:Keywords:	immunoglobulin		
F:20-85/Domain:	immunoglobulin homology <IMM>		
Query Match	50.8%	Score	1156; DB 2; Length 377;
Best Local Similarity	74.3%	Pred. No.	3.6e-64;
Matches	223;	Conservative	16; Mismatches 37; Indels 24; Gaps
Qy	110	VOAGTPERNT--VCKRCPDGFFSNETSskap-CRKHTNCVSFGLLLTQGNATHDNCISG	406
Db	98	VELKTPLGDTHTTCPRCEP-KSGDTPPCPRCEPKSCD-----	136
Qy	167	NSESTQKGDIDTVDKTHTCTPCPAPELLGGPSVFLFPPPKPDTLMISRTPEVTCVVVDV	236
Db	137	TPPCPRCEPKSCDTPPCPRCPAPELLGGPSVFLFPPPKPDTLMISRTPEVTCVVVDV	196
Qy	227	SHEPDEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKKYSNK	286
Db	197	SHEPDEVQFKWYVDGVEVHNAKTKPREEQYNSTFRVYSLTVLHQDWLNGKEYCKKYSNK	256
Qy	287	ALPAPIETISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ	346
Db	257	ALPAPIETISKTKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ	316
Qy	347	PENNYKTPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHREALHNHYTQKSLSLSPG	406
Db	317	PENNYNTTPMLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHREALHNHYTQKSLSLSPG	376
RESULT	7		
A60764			
Ig gamma-3 chain C region, form LAT - human			
C:Species:	Homo sapiens (man)		
C:Date:	14-May-1993	#sequence_revision	14-May-1993
C:Accession:	A60764	#text_change	16-Jul-1991
R:Huck, S.; Lefranc, G.; Lefranc, M.P.			
Immunogenetics	30, 250-257, 1989		
A:Title:	A human immunoglobulin IGHG3 allele (Gmb0, bl, c3, c5, u) with an		
A:Reference number:	A60764; MUID:90007613		
A:Accession:	A60764		
A>Status:	Preliminary		

A:Molecule type: DNA  
A:Residues: 1-377 <HUC>

C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: immunoglobulin

F:20-85/Domain: immunoglobulin homology <IMX>

Query Match 50.7%; Score 1154; DB 2; Length 377;  
Best Local Similarity 74.3%; Pred. No. 4.8e-64;  
Matches 223; Conservative 16; Mismatches 37; Indels 24; Gaps 4;

QY 110 VQATPERNT--VCKRCPDGFSSNETSKAP-CRKHNCSVFGLLLTQKGNATHDNTCSG 166

Db 98 VELTPLGDTHTTCRCPEP-KSCDTPPCRCPEPKSCD----- 136

QY 167 NSESTQKGDVTDKTHKCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDV 226

Db 137 TTPCPCRCPEPKSCDTPPCRCPEPKKDTLMISRTPEVTCVVVDV 196

QY 227 SHEDPEVKFNMYVDGVEVHNAKTPREEQYNSTYRVSVLTFLVLDHQLNGKEYCKKYSNK 286

Db 197 SHEDPEVQKMYVDGVEVHNAKTPREEQYNSTYRVSVLTFLVLDHQLNGKEYCKKYSNK 256

QY 287 ALPAPIEKTISKAGQPREPOVYILPSPRDELTKNQVSLTCLVKGFPSPDIATWESNGQ 346

Db 257 ALPAPIEKTISKAGQPREPOVYILPSPRDELTKNQVSLTCLVKGFPSPDIATWESNGQ 316

QY 347 PENNYKTPPPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPG 406

Db 317 PENNYKTPPPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPG 376

## RESULT 8

G2HU

Ig gamma-2 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 30-Apr-1981 #sequence\_revision 13-Jun-1983 #text\_change 21-Jul-2000

C:Accession: A93906; A92809; A90752; A93132; A02148

Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982

A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con

A:Reference number: A93906; MUID:82197621

A:Accession: A93906

A:Molecule type: DNA

A:Residues: 1-326 &lt;ELL&gt;

A:Cross-references: GB:V00554; GB:J00230; NID:G32759; PIDN:CAB58438.1; PID:G6066056

A:Note: Lys-326 is probably removed posttranslationally

R:Wang, A.C.; Tung, E.; Fudenberg, H.H.

J. Immunol. 125, 1048-1054, 1980

A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f

A:Reference number: A92809; MUID:81007873

A:Contents: myeloma protein Til

A:Accession: A92809

A:Molecule type: protein

A:Residues: 1-19,'Q',21-57,'Z',59,'A',61-193,'D',195-325 &lt;WAN&gt;

A:Note: Trp-156 is at or near the complement-binding site

R:Connell, G.E.; Parr, D.M.; Hofmann, T.

Can. J. Biochem. 57, 758-767, 1979

A:Title: The amino acid sequences of the three heavy chain constant region domains of a

A:Reference number: A90752; MUID:80001357

A:Contents: myeloma protein zie

A:Accession: A90752

A:Molecule type: protein

A:Residues: 1-24,'E',26-57,'EV',60-85;132-171,'ZZZ',175,'B',177-193,'D',195-196,'Q',198

A:Note: this sequence has since been revised

R:Hofmann, T.; Parr, D.M.

Mol. Immunol. 16, 923-925, 1979

A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g

A:Reference number: A93132; MUID:80114419

A:Contents: Zie

A:Accession: A93132

A:Molecule type: protein

A:Residues: 238-275 &lt;HOF&gt;

R:Hofmann, T.; Parr, D.M.

Submitted to the Atlas, March 1980

A:Reference number: A94591

A:Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268

A:Note: the revised sequence differs from that shown in having 60-Ala and in the amid

ned

R:Milstein, C.; Frangione, B.

Biochem. J. 121, 217-225, 1971

A:Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.

A:Reference number: A90253; MUID:72033500

A:Contents: annotation; myeloma protein Sa, disulfide bonds

R:Frangione, B.; Milstein, C.; Pink, J.R.L.

Nature 221, 145-148, 1969

A:Title: Structural studies of immunoglobulin G.

A:Reference number: A93157; MUID:69064124

A:Contents: annotation; Sa, disulfide bonds

C:Genetics:

A:Gene: GDB:IGHG2

A:Cross-references: GDB:119338; OMIM:147110

A:Map position: 14q32.33-14q32.33

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into

C;Superfamily: immunoglobulin C region; immunoglobulin homology

C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:20-85/Domain: immunoglobulin homology &lt;IMI&gt;

F:239-306/Domain: immunoglobulin homology &lt;IM2&gt;

F:14/Disulfide bonds: interchain (to light chain) #status experimental

F:27-83;140-200;246-304/Disulfide bonds: #status experimental

F:102;103;106;109/Disulfide bonds: interchain (to heavy chain) #status experimental

F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.5%;

Best Local Similarity 87.0%; Pred. No. 9.6e-64;

Matches 215; Conservative 9; Mismatches 11; Indels 12; Gaps 3;

QY 171 TQKGID----VYVDKT-----HTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEV 219

Db 80 TYTCNVDPKPSNTKVDKTVKCCVCCPCAPP-VAGPSVFLPPKPKDTLMISRTPEV 138

QY 220 TCVVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEQYNSTYRVSVLTFLVLDHQLNGKEY 279

Db 139 TCVVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEQYNSTYRVSVLTFLVLDHQLNGKEY 198

QY 280 KCKVSNKALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFPSPDIATV 339

Db 199 KCKVSNKALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFPSPDIATV 258

QY 340 EWESNGQPENNYKTPPPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQ 399

Db 259 EWESNGQPENNYKTPPPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQ 318

QY 400 SLSLSPG 406

Db 319 SLSLSPG 325

## RESULT 9

G4HU

Ig gamma-4 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 16-Jul-1999

C:Accession: A90933; A90249; A02150

R:Ellison, J.; Buxbaum, J.; Hood, L.

DNA 1, 11-18, 1981

A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.

A:Reference number: A90933; MUID:83157104

A:Accession: A90933

A:Molecule type: DNA

A:Residues: 1-327 &lt;ELL&gt;

A:Note: the sequence was determined from the germline gene

R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.





[illegible]

Query Match 39.9%; Score 907; DB 2; Length 328;  
Best Local Similarity 65.9%; Pred. NO. 6.1e-49;  
Matches 170; Conservative 33; Mismatches 35; Indels

	Query Match	39.9%;	Score 907;	DB 2;	Length 328;
	Best Local Similarity	65.9%;	Pred. No. 6.1e-49;		
	Matches 170;	Conservative 33;	Mismatches 35;	Indels 20;	Gaps
QY	165	SGNSESTQKGID-----VTVDK---THTCPPCP-----APELLGGPSVFLFPPPKPT 210           :           :           :           :           :           :           :           :           :           :           :           :			
Db	74	SLSKS SVTCNVNHPATTTKVDRVRGTKTPCPCIPACESP----GPSVFIEFPKPCKT 129           :           :           :           :           :           :           :           :           :           :           :           :			
QY	211	LMSRTEPVTCVVVDVSHEDPEVKFNWYVDGVGEVHNAAKTPREEQYNSTRYVVSMTLVH 270           :           :           :           :           :           :           :           :           :           :           :           :           :           :			
Db	130	LMSRTPQVTCVVVDVSDENPEVQFSWYVDGVGEVHTAQTARKKEQSFTSYRVYSVLPIQH 189           :           :           :           :           :           :           :           :           :           :           :           :           :           :			
QY	271	QDWLNGKEYECKSKNALPAPIETKTISKAKGQPREFQVITPLSPSRDELTKNQSVLECLVK 330           :           :           :           :           :           :           :           :           :           :           :           :           :           :			
Db	190	QDWLNGKEYECKVNNDLPAPITRIISKAGQTRFQVITPLPHAEELSRKSVISICLI 249           :           :           :           :           :           :           :           :           :           :           :           :           :           :			
QY	331	GFYPSTDIAVESWGQ--PENNYKTTTPPVLDSDGSFELYSKLTVDKSRWQQGNVFCSVM 388           :           :           :           :           :           :           :           :           :           :           :           :           :           :			
Db	250	GFYPDIDVQRNGQDEPEGNRYTTPPOQDVDTGYFLYSKFSVDKASWGGGIFOCAYM 309           :           :           :           :           :           :           :           :           :           :           :           :           :           :			
QY	389	HEALHHNYTKSLSPG 406           :           :           :           :           :           :           :           :           :           :           :           :           :           :			
Db	310	HEALHHNYTKSISKTPG 327           :           :           :           :           :           :           :           :           :           :           :           :           :           :			

Query Match	39.9%;	Score 907;	DB 2;	Length 328;
Best Local Similarity	65.9%;	Pred. NO. 6.1e-49;		
Matches 170; Conservative	33;	Mismatches 35;	Indels 20;	Gaps 5;



Db 221 KGAPRMPDVYTLPPSRDELSKSVSVTCLIIINFPFADIHVEWASNRVPVSEKEYKNTPEI 280  
Qy 358 LDSGSGFFLYSKLTVDKSRWQQGNVFCVMHEALHNHYTOKSLSLSPG 406  
Db 281 EDADGSFELYSKLTVDKSAWDQGTYYTCSVMHEALHNHYTOKAISRSFG 329

Search completed: March 1, 2001, 09:15:48  
Job time: 146 sec

Result No.	Score		Query Match		Length	DB	ID	Description
	Score	Match						
1	1490	65.5	485	13	R24016	Fusion protein TNF		
2	1487	65.4	518	15	R51003	Sequence of a reco		
3	1349	59.3	376	19	W60337	Antigenic peptide		
4	1334	58.7	376	18	W50287	Human Fas antigen		
5	1330	58.5	438	16	R81882	Plasmid pDC406/Ox4		
6	1330	58.5	438	19	W48976	Ox40/FC protein. C		
7	1283	56.4	764	21	W68949	Fusion protein of		
8	1257.5	55.3	664	19	W71603	Human neuturin re		
9	1257.5	55.3	664	21	X80123	Human NTNR alpha		
10	1254.5	55.2	664	19	W71604	Rat neuturin rece		
11	1254.5	55.2	664	21	X80124	Rat NTNR alpha and		
12	1253	55.1	963	19	W70540	Integrin beta-1 ch		



```
|||||
Db 368 nstyrvvsvltvqhwdngdykckvsnkalpampgktiskakggpreqvylpsrd 427
QY 317 ELTKNOVSLTCLVKGFPSDIAVWESGOPENNYKTPPVLDSDGSFFLYSLKLTVDKSR 376
Db 428 eltknqvsclclvkgyfprhriavewesngpennykttppvlsgdsfflyslkltvdksr 487
QY 377 WQGNVFCSCVMHEALHNHYTKSLSPG 406
Db 488 wqggnvfscvmhealhnhytqkslsisp 517

RESULT 3
ID W60037 standard; Protein; 376 AA.
XX
AC W60037;
XX
DT 11-SEP-1998 (first entry).
XX
DE Antigenic peptide hFas (nd29) containing Fc region.
XX
KW Fas ligand; Fas antagonist; apoptosis related disease; liver disease;
KW heart failure; kidney failure; graft-versus-host disease; antibody;
KW myocardial infarction; ischemic restenosis; endotoxemic shock.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT Protein
FT /note="hFas antigen signal peptide"
FT /note="hFas (nd29) protein"
XX
PN WQ9818487-AL.
XX
PD 07-MAY-1998.
XX
PF 31-OCT-1997; 97WO-JP03978.
XX
PR 26-SEP-1997; 97JP-0262521.
PR 31-OCT-1996; 96JP-0290459.
PR 27-DEC-1996; 96JP-0351718.
XX
PA (MOCH ) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSAKA BIOSCIENCE INST.
XX
PI Nagata S, Suda T, Yatomi T;
XX
DR WPI; 1998-271925/24.
DR N-PSDB; V34430.
XX
PT Use of Fas antagonist for treatment and prevention of
PT apoptosis-related diseases - such as heart or kidney failure,
PT graft-versus-host disease or liver disease
XX
PS Examples; Fig 5-9; 86pp; Japanese.
XX
CC This represents the antigenic peptide hFas (nd29) containing the Fc
CC region. The invention provides the use of Fas antagonist as an agent for
CC the treatment and prevention of apoptosis-related diseases. The Fas
CC antagonist can be a partial Fas antigen peptide containing the
CC extracellular part of the protein, but lacking the signal sequence, an
CC anti-Fas antibody, or an anti-Fas ligand antibody, where the antibody is
CC preferably a humanized antibody. The Fas antagonist is used in the
CC treatment and prevention of diseases such as myocardial infarction, heart
CC failure, ischemic heart disease, acute kidney failure, graft-versus-host
CC disease, ischemic restenosis of the heart, liver or kidney, and
CC endotoxemic shock, and also as an organ preservative in transplantation.
CC The agent is of low toxicity but effectively inhibits the Fas/Fas ligand
CC system.
XX
SQ Sequence 376 AA;
```

```
Query Match 59.3%; Score 1349; DB 19; Length 376;
Best Local Similarity 66.3%; Pred. No. 1.5e-76;
Matches 270; Conservative 15; Mismatches 60; Indels 62; Gaps 9;
QY 8 LHDEETSHQLLCKCPPTGLYKQHCTAKW-KTVCAFCPD-HYYTDSMHTSDECLYCSVP 65
Db 23 lhhdggfchk----pcppgerkardctvngdepdcvpcqegkeytdkafhsskrrcr-l 77
QY 66 CKELQYVVKQE--CNRTNRRVCECKEGRYLEIEFCLKHRSCLPP-----GFGVVQAGTPERT 119
Db 78 cdeghgleveinctrtqtkcrckpnfnstvc---ehcdpctkcehgfiikeetltsnt 134
QY 120 VCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTKGNATHDNCISGNSESTQKCGIDVT 179
Db 135 kcke--egsrsnepks----- 148
QY 180 VDKTHCTPCPAPPELLGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMY 239
Db 149 cdkthtccppcappelggpsvflfppkpkdtlmisrtpevtcvvvdvshedpevfwnyv 208
QY 240 DGEVEHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISK 299
Db 209 dgvevhnaktpreegynstyrvvsvltvqhwdngkeykckvsnkalpapietiska 268
QY 300 KGQPREPQVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVWESGOPENNYKTPPVLD 359
Db 269 kgqprepqvylpsrdeltnkvsltcclvkgyfypsdiavwesngpennykttppvld 328
QY 360 SDGSFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPG 406
Db 329 sdgsfflyslkltvdksrwqggnvfscvmhealhnhytqkslsisp 375

RESULT 4
W50287
ID W50287 standard; Protein; 376 AA.
XX
AC W50287;
XX
DT 16-JUL-1998 (first entry)
XX
DE Human Fas antigen derivative/IgG1 Fc fusion.
XX
KW Human; Fas antigen; derivative; apoptosis regulation; gene therapy;
KW treatment; diabetes; arthritis; lupus; hepatitis; influenza; HIV;
KW apoptosis modulation; immunoglobulin G1 Fc; IgG1 Fc; fusion.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT /label= sig_peptide
FT /label= mat_peptide
XX
PN WQ9742319-AL.
XX
PD 13-NOV-1997.
XX
PF 01-MAY-1997; 97WO-JP01502.
XX
PR 02-MAY-1996; 96JP-0135760.
XX
PA (MOCH ) MOCHIDA PHARM CO LTD.
PA (OSAB-) OSAKA BIOSCIENCE INST.
XX
PI Nagata S, Nakamura N;
XX
DR WPI; 1997-558981/51.
DR N-PSDB; V07004.
XX
```

PT Fas antigen derivative containing modified extracellular region -  
 PT has low antigenicity, promotes apoptosis and is useful in treatment  
 of viral and other diseases  
 XX  
 PS Disclosure; Fig 4; 102pp; Japanese.  
 XX  
 CC The present sequence is a Fas antigen derivative/IgG1 Fc  
 CC fusion, which contains a Fas antigen extracellular region lacking  
 CC one or more amino acid residues in the region from the  
 CC amino-terminal to (but excluding) the 1st cysteine residue  
 CC (preferably at least 29 residues are deleted).  
 CC The derivative is an effective regulator of apoptosis and can be  
 CC used (either by administration of the polypeptide, or by the use  
 CC of the coding DNA in gene therapy) to treat a range of diseases,  
 CC e.g. diabetes, arthritis, lupus and in particular viral diseases  
 CC such as hepatitis, influenza and HIV, by modulating apoptosis of  
 CC virus-infected cells.  
 XX  
 SQ Sequence 376 AA;

Query Match 58.7%; Score 1334; DB 18; Length 376;  
 Best Local Similarity 62.6%; Pred. No. 1.3e-75;  
 Matches 267; Conservative 16; Mismatches 62; Indels 62; Gaps 9;

QY 8 LHYDEETSHQLLCKDPCPGTYLKQHTAKW-KTVCAPCPD-HYITDSWHTSDECLYCSVPV 65  
 DB 23 lhhdgqfchh---pcppgkerkdctvngdepdcvpcqegkeytdkafsfkrrcr-l 77  
 QY 66 CKEIYQVKE--CNRTHNRVCEKGRYLEIEFCLKHSRCP--GFGVVOAGTPERNT 119  
 DB 78 cdegghleincrtqtkrcpfnfstvc---ehcdpctkcehglikectltst 134  
 QY 120 VKRCPCDFFSNESKAPCRKHTNCVFGLLLTQKGNATHDNCSGNSESTQKCGIDVT 179  
 DB 135 kcke--egdsrnepsk----- 148  
 QY 180 VDKHTCCPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTVVDSHEDPEVKFNMYV 239  
 DB 149 cdktctcpcpapellgspvflfppkpkdtlmsptpevtcvvdsvedevkfnmyv 208  
 QY 240 DGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIKA 299  
 DB 209 dgvevhnaktprreeqynstyrvvsvltvlhqdwingkeyckvsnkalpapiektiska 268  
 QY 300 KGOPEPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 359  
 DB 269 kgqprepqvylppsrdeitknqvsitclvkgyfypsdiavewesngqpennykttppvld 328  
 QY 360 SDGSFFLYSLTKVDSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 406  
 DB 329 sdgsfflyslntvdksrwqqgnvfscsvmhealhnhytqkslsisp 375

RESULT 5  
 R81882  
 ID R81882 standard; Protein; 438 AA.  
 AC R81882;  
 XX  
 DT 30-MAR-1996 (first entry)  
 XX  
 DE Plasmid pDC406/OX40/Fc\* encoding an OX40/Fc mutein protein.  
 XX  
 KW OX40; OX40-L; cytokine; cell surface molecule; plasmid;  
 KW pDC406/OX40/Fc\*; membrane glycoprotein.  
 XX  
 OS Synthetic.  
 XX  
 PN US5457035-A.  
 XX  
 PD 10-OCT-1995.  
 XX

PF 23-JUL-1993; 93US-0097827.  
 PR 23-JUL-1993; 93US-0097827.  
 XX  
 PA (IMMV ) IMMUNEX CORP.  
 XX  
 PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;  
 DR WPI: 1995-357992/46.  
 DR N-PSDB: T00829.  
 XX  
 PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors  
 PT and host cells, used to produce recombinant ligand used in e.g.  
 PT prim. T cell culture, to modulate immune response etc.  
 XX  
 PS Example 2; Column 35-38; 26pp; English.  
 XX  
 CC This plasmid encodes an OX40/Fc antibody fragment mutein protein,  
 CC and is used to express a soluble OX40/Fc mutein fusion protein for  
 CC use in detecting cDNA clones encoding a OX40 ligand. The Fc  
 CC fragment may be derived from human IgG1, and the plasmid may be  
 CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell  
 CC line. Culture supernatant was purified by affinity chromatography  
 CC and this was used, together with labeled goat anti-human IgG to  
 CC screen various cell lines.  
 XX  
 SQ Sequence 438 AA;

Query Match 58.5%; Score 1330; DB 16; Length 438;  
 Best Local Similarity 62.6%; Pred. No. 2.7e-75;  
 Matches 263; Conservative 23; Mismatches 80; Indels 54; Gaps 6;

QY 20 CDKCPPTYLKQHTAKWKTVCAPCPDHYITD--SWHTSDECLYCSVPCKELQYVQSCN 77  
 DB 39 ccreqpgghgmvrchdtrdtlchpcetgfneavnydtckqctqcnh--rsgselkgnct 96  
 QY 78 RTHNRVCEKGRYLEIEFCLKHSRCPGFGVVOAGTPERNTVCKRCPDGPFNSNETSKA 137  
 DB 97 ptqdtvcr-----pgtprqdsqykgldcvpcppghfs--pgnq 137  
 QY 138 PCRHTNGSVFGLLLTQKGNATHDNCSGNS-----ESTQKCGIDVT----- 180  
 DB 138 ackptwtctslgkqtrhpasdsldavcedrsilatllwetqrptfrptvtggtwprts 197  
 QY 181 -----DKHTCCPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTVVDSV 226  
 DB 198 elpatptliveprscdkthtccpapeagapsvflfppkpkdtlmsrtpevtcvvvdv 257  
 QY 227 SHEDPEVKFNMYVDGVEVHNKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNK 286  
 DB 258 shedpevkfnmyvdgvevhnaktprreeqynstyrvvsvltvlhqdwingkeyckvsnk 317  
 QY 287 ALPAPIEKTIKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ 346  
 DB 318 alpapiektiskakgqprepqvylppsrdeitknqvsitclvkgyfypsdiavewesngq 377  
 QY 347 PENNYKTTTPPVLDSDGSFFLYSLTKVDSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 406  
 DB 378 pennykttppvldsdgsfflyskitvdksrwqqgnvfscsvmhealhnhytqkslsisp 437

RESULT 6  
 W48976  
 ID W48976 standard; Protein; 438 AA.  
 AC W48976;  
 XX  
 DT 25-SEP-1998 (first entry)  
 XX  
 DE OX40/Fc mutein.  
 XX  
 KW OX40/Fc; cytokine; T cell antigen; TH-2 immune response; OX40-L;



[illegible]

Qy	227	SHDEPVKFNWYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKCVSNK	286
Db	258	shedpevkfnwyvdgvevhnaktkpreegynstyrsvsltlvhhgdwlogkeyckcvsnk	317
Qy	287	ALPAPTEKTSKAGGPREQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVEWESNGQ	346
Db	318	alpaptektiskaggrepqvylppsrdelckngvslclvkgfypsdiavewesngq	377
Qy	347	PENNYKTPPVLSDSGSFFLYSLKTLVDKSRWQOGNFGSCVMHEALHNHYTKQSLSLSPG	406
Db	378	pennyktppldsdgsfflysktlvdkserwqgnfgscvmhealhnhytkksls	437
RESULT	7		
Y68949			
ID	Y68949	standard; Protein; 764 AA.	
XX			
XX	Y68949;		
XX			
DT	30-MAY-2000	(first entry)	
XX			
DE		Fusion protein of murine delta-related protein and human IgG Fc.	
XX			
KW		Cell development cycle; Delta family; membrane surface-bound ligand;	
KW		endothelial cell biology; gene therapy; subcortical infarct;	
KW		cerebral autosomal dominant arteriopathy; leucoencephalopathy;	
KW		ischemic stroke; chimera.	
XX			
OS		Chimeric - Mus sp.	
OS		Chimeric - Homo sapiens.	
XX			
Key		Location/Qualifiers	
FT	Protein	1.529	
FT		/note= "extracellular region of the murine	
FT		delta-related protein"	
FT	Protein	533..764	
FT		/note= "human IgG Fc portion"	
XX			
PN	W0200006726-A2.		
XX			
PD	10-FEB-2000.		
XX			
PF	12-JUL-1999;	99WO-US15710.	
XX			
PR	27-JUL-1998;	98US-0123168.	
XX			
PA	(AMGE-) AMGEN INC.		
PI	Shutter JR, Stark KL;		
XX			
DR	WPI; 2000-195294/17.		
DR	N-PSDB; Z60926.		
XX			
PT		Cell development cycle protein of delta family useful for treating	
PT		various disorders associated with central nervous system e.g. cerebral	
PT		autosomal dominant arteriopathy and ischemic strokes	
XX			
PS	Example 6; Page 169-171; 171pp; English.		
XX			
CC		The present sequence represents a fusion protien of the extracellular	
CC		domain of a murine polypeptide, which a member of the cell	
CC		development cycle protein family known as the Delta family of	
CC		mammalian membrane surface-bound ligands, and the human immunoglobulin	
CC		G (IgG) Fc portion. The murine delta-related protein gene is expressed	
CC		within vascular endothelium indicates a role for the polypeptides	
CC		in the control of endothelial cell biology. The murine polynucleotide	
CC		was identified from a white adipose tissue cDNA library. The polypeptide	
CC		is useful for identifying receptors, which bind to and/or are activated	
CC		by the polypeptide. The polynucleotide is useful in gene therapy of	
CC		cerebral autosomal dominant arteriopathy with subcortical infarcts and	
CC		leucoencephalopathy, an autosomal dominant disorder causing ischemic	
CC		strokes.	



```

XX 19-MAY-2000 (first entry)
XX Human NTNR alpha and IgG fusion protein SEQ ID NO:16.
DE Human neurturin receptor alpha; NTNR alpha; splenic haematopoiesis;
KW anaemia; thrombocytopaenia; hypoplasia; haemorrhage.
XX Homo sapiens.
OS Synthetic.
XX US6025157-A.
XX 15-FEB-2000.
XX 24-OCT-1997; 97US-0957063.
XX 18-FEB-1997; 97US-0038839.
XX 09-JUN-1997; 97US-0049818.
XX (GETH ) GENENTECH INC.
XX Hynes MA, Rosenthal A, Klein RD;
XX WPI; 2000-181808/16.
XX N-PSDB; 291459.
XX Isolated nucleic acid molecule encodes a neurturin receptor-alpha amino
XX acid sequence excluding the N-terminal signal peptide -
XX Claim 3; Column 91-96; 78pp; English.
XX The present invention describes a neurturin receptor alpha (NTNR alpha).
XX NTNR alpha binds neurturin. The NTNR alpha nucleic acid molecule is
XX useful for the expression of NTNR alpha, which may be used to identify
XX agonist and antagonist compounds having therapeutic applications, such
XX as enhancing splenic haematopoiesis, treating anaemia,
XX thrombocytopaenia, hypoplasia, or haemorrhage. The present sequence
XX is a NTNR alpha and IgG (immunoglobulin G) fusion protein from the
XX present invention.
XX Sequence 664 AA;

Query Match 55.3%; Score 1257.5; DB 21; Length 664;
Best Local Similarity 62.7%; Pred. No. 1.3e-70;
Matches 264; Conservative 25; Mismatches 55; Indels 77; Gaps 14;

QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDCLYCSPVKELQV 72
DB 274 ncrasyqtvtst-cpadnyqacilgsvagmigmfntpnvdsstgt---ivvspwc----- 323
QY 73 KQECNTHNRVCEKRGYLEIEF---CLKH-----RSCP--PGFGVVOAGTP 115
DB 324 --scrgsgnmeece--kflr-dftenpcrlnaigafngtdvvnvspkqpsfqatqprv 378
QY 116 ERNTVCKRCPDFFNETSSKAPCRKHTNCVSFGLLLTQGNATHNICSNGSESTQKCG 175
DB 379 ext---pslpdd-lsdstslgts-----vittctsvqeqglkannskelsmcf 422
QY 176 IDVT-----VDKTHCPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVD 225
DB 423 teltniipgrdpvdkthtccpapelilggpsvflfpkpkdtlmisrtpevtcvvvd 482
QY 226 VSHDEPEVKFNVDGVVHNAKTRPREQYNSTYRVSVLTVLHWDLNGKEYKCKVSN 285
DB 483 vshedpevkfnvvdgvevhnaktprreeqynstyrsvsvltvlhqdwlngkeykckvsn 542
QY 286 KALPAPIETKISAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNG 345
DB 543 kalpapietkiskagqprepqvtytlppsrdeentknqsvltclvkgyfypsdiavehesng 602
QY 346 QPENNYKTPPVLDSDGSEFFLYSKLTVDKSRMQQGNVFCSCVMHEALHNHYTKSLSLSP 405

```

```

Db 603 qpennyktppvldsdgsefflyskltdvksrwqgnvfcscvmhealhnhytkslslsp 662
QY 406 G 406
DB 663 g 663

RESULT 10
W71604
ID W71604 standard; Protein; 664 AA.
XX AC W71604;
XX 19-NOV-1998 (first entry)
XX DE Rat neurturin receptor alpha/Fc sequence (Iff2a) fusion protein.
XX KW Human; neurturin receptor alpha; NTNR-alpha; variant; chimeric;
XX KW fusion protein; immunoadhesion; ret-expressing cell; neurological;
XX KW renal; haematological disease.
XX OS Synthetic.
XX OS Homo sapiens.
XX OS Rattus sp.
XX PN WO9836072-A1.
XX 20-AUG-1998.
XX 17-FEB-1998; 98WO-US03179.
XX 24-OCT-1997; 97US-0957063.
XX 18-FEB-1997; 97US-0802805.
XX 09-JUN-1997; 97US-0871913.
XX (GETH ) GENENTECH INC.
XX Hynes MA, Klein RD, Rosenthal A;
XX WPI; 1998-467175/40.
XX N-PSDB; V58007.
XX New polypeptide(s) based on human neurturin receptor alpha and
XX related nucleic acid - useful for increasing survival of
XX ret-expressing cells for treating e.g. neurological, renal and
XX haematological diseases
XX Example 3; Page 84-86; 116pp; English.
XX The present sequence represents rat neurturin receptor alpha
XX (NTNR-alpha)/Fc sequence (Iff2a) fusion protein, from an example of the
XX present invention. NTNR-alpha proteins can be used: (a) to identify
XX molecules that bind specifically to it (potential agonists and
XX antagonists) and to purify such compounds; (b) to modulate response of
XX cells to neurturin (NTN); (c) to increase survival of ret-expressing
XX cells or to activate Ret on the surface of cells (soluble glial derived
XX neurotrophic factor receptor (GDNFR alpha) may be used the same way);
XX (d) to increase the half-life of cognate ligands (especially NTN); (e)
XX diagnostically to determine serum levels of its ligands; and (f) as
XX animal feed additive or molecular weight marker. NTNR-alpha, its genes,
XX (ant)agonists and antisense nucleic acids, are useful in vivo or ex vivo
XX for treating conditions related to abnormal NTN activity or response,
XX particularly neurological (central or enteric), renal or haematopoietic
XX (spleen) diseases or injuries. Ab may be agonists or antagonists for
XX therapeutic use (e.g. as antagonists to treat excessive/unwanted NTNR
XX alpha expression, e.g. in some tumours), also reagents for immunoassay
XX and affinity purification.
XX Sequence 664 AA;

Query Match 55.2%; Score 1254.5; DB 19; Length 664;

```

Best Local Similarity 62.8%; Pred. No. 2e-70; Matches 262; Conservative 30; Mismatches 56; Indels 69; Gaps 13;

QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDECLYSPVCKELQYV 72  
DB 274 ncrasyrtits-cpadnyqacqsyagmigdmtpnvydsnptg---ivspwc----- 323  
QY 73 KQECNRTHNRVCECKEGRYLEIEF---CLKHRSPPGFGVVOAGTPEPNTVCKRCPDGF 128  
DB 324 --ncrgsgnmeeeee--kflr-dftenpcl--rnaiaqfg-----ngtdvnmsspg- 367  
QY 129 FSNETSskapcrkht-----NCSVFGLLLTOKGNATHDNICSGNSESTQKCGIDVT 179  
DB 368 -psipatqarvcktpelpddlsdstslgtsvittctsiqegqikannskelsmcfelt 426  
QY 180 -----VDKTHTCPPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHE 229  
DB 427 tnispgsrpdkvcktpcpapellggpsvflfpkpkdtlmisrtpevtcvvdvshe 486  
QY 230 DPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP 289  
DB 487 dpevkfnwydgvgevhnaaktpreeqynstyrvvsvltvlhqdwlngkeyckvsnkalp 546  
QY 290 APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPEN 349  
DB 547 apiektiskakgqprepqvylppsrreemtknqvsllclvkgyfypsdiawesngqpen 606  
QY 350 NYKTPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSMHEALHNNHYTKSLSLSPG 406  
DB 607 nykttppvldsdgsfflyskltvdksrwqggnvfscsmhealhnhytqkslsisp 663

RESULT 11

Y80124  
ID Y80124 standard; Protein; 664 AA.  
XX AC Y80124;  
XX DT 19-MAY-2000 (first entry)  
XX DE Rat NTNR alpha and IgG fusion protein SEQ ID NO:18.  
XX KW Human; neuturin receptor alpha; NTNR alpha; splenic haematopoiesis;  
XX KW anaemia; thrombocytopaenia; hypoplasia; haemorrhage.  
XX OS Rattus sp.  
XX OS Synthetic.  
XX PN US6025157-A.  
XX PD 15-FEB-2000.  
XX PF 24-OCT-1997; 97US-0957063.  
XX PR 18-FEB-1997; 97US-0038839.  
XX PR 09-JUN-1997; 97US-0049818.  
XX PA (GETH ) GENENTECH INC.  
XX PI Hynes MA, Rosenthal A, Klein RD;  
XX DR WPI; 2000-181808/16.  
XX DR N-PSDB; 291460.  
XX PT Isolated nucleic acid molecule encodes a neuturin receptor-alpha amino  
XX PT acid sequence excluding the N-terminal signal peptide -  
XX PS Claim 3; Column 97-102; 78pp; English.  
XX CC The present invention describes a neuturin receptor alpha (NTNR alpha).  
XX CC NTNR alpha binds neuturin. The NTNR alpha nucleic acid molecule is  
XX CC useful for the expression of NTNR alpha, which may be used to identify  
XX CC agonist and antagonist compounds having therapeutic applications, such

CC as enhancing splenic haematopoiesis, treating anaemia,  
CC thrombocytopaenia, hypoplasia, or haemorrhage. The present sequence  
CC is a NTNR alpha and IgG (immunoglobulin G) fusion protein from the  
CC present invention.  
XX SQ Sequence 664 AA;

Query Match 55.2%; Score 1254.5; DB 21; Length 664;

Best Local Similarity 62.8%; Pred. No. 2e-70; Matches 262; Conservative 30; Mismatches 56; Indels 69; Gaps 13;

QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDECLYSPVCKELQYV 72  
DB 274 ncrasyrtits-cpadnyqacqsyagmigdmtpnvydsnptg---ivspwc----- 323  
QY 73 KQECNRTHNRVCECKEGRYLEIEF---CLKHRSPPGFGVVOAGTPEPNTVCKRCPDGF 128  
DB 324 --ncrgsgnmeeeee--kflr-dftenpcl--rnaiaqfg-----ngtdvnmsspg- 367  
QY 129 FSNETSskapcrkht-----NCSVFGLLLTOKGNATHDNICSGNSESTQKCGIDVT 179  
DB 368 -psipatqarvcktpelpddlsdstslgtsvittctsiqegqikannskelsmcfelt 426  
QY 180 -----VDKTHTCPPCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHE 229  
DB 427 tnispgsrpdkvcktpcpapellggpsvflfpkpkdtlmisrtpevtcvvdvshe 486  
QY 230 DPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALP 289  
DB 487 dpevkfnwydgvgevhnaaktpreeqynstyrvvsvltvlhqdwlngkeyckvsnkalp 546  
QY 290 APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPEN 349  
DB 547 apiektiskakgqprepqvylppsrreemtknqvsllclvkgyfypsdiawesngqpen 606  
QY 350 NYKTPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSMHEALHNNHYTKSLSLSPG 406  
DB 607 nykttppvldsdgsfflyskltvdksrwqggnvfscsmhealhnhytqkslsisp 663

RESULT 12

W70540  
ID W70540 standard; Protein; 963 AA.  
XX AC W70540;  
XX DT 26-JAN-1999 (first entry)  
XX DE Integrin beta-1 chain.  
XX KW Integrin; beta-1 chain; immunoglobulin; chimeric; heterodimer complex;  
XX KW inhibitor; binding; ligand; blood platelet; hemostatic; diagnostic agent;  
XX KW human.  
XX OS Homo sapiens.  
XX PH Key Location/Qualifiers  
XX FT Peptide 1..20  
XX FT /note= "signal peptide"  
XX FT Protein 21..963  
XX FT /note= "mature protein"  
XX PN W09832771-A1..  
XX PD 30-JUL-1998.  
XX PF 29-JAN-1998; 98WO-JP00370.  
XX PF 29-AUG-1997; 97JP-0234544.  
XX PR 29-JAN-1997; 97JP-0015118.  
XX PA (TORA ) TORAY IND INC.

XX Kainoh M, Tanaka T;  
 XX WPI: 1998-427881/36.  
 DR N-PSDB; V33773.  
 XX Integrin-immunoglobulin chimeric protein heterodimer complexes as  
 PT platelet substitutes - contain the alpha and beta integrin chains  
 PT associated in stable state and bind to extracellular matrix in the  
 PT presence of plasma components  
 XX  
 PS Claim 9; Pages 50-57; 87pp; Japanese.  
 XX This represents an integrin beta-1 chain. The invention provides  
 CC integrin-immunoglobulin chimeric protein heterodimer complexes that  
 CC comprise an integrin alpha or beta chain associated with an  
 CC immunoglobulin light or heavy chain. These chimeric proteins form  
 CC heterodimer complexes, in particular with a chimeric protein containing  
 CC an integrin alpha chain and an immunoglobulin chain with a chimeric  
 CC protein containing an integrin beta chain and an immunoglobulin chain;  
 CC the immunoglobulin chain in each case may be a heavy chain, or one of the  
 CC two may be a light chain. The integrin alpha chain is preferably alpha 4  
 CC or alpha 2 and the integrin beta chain is preferably beta 1. Animal cells  
 CC transformed with vectors containing the DNA coding for the above chimeric  
 CC proteins can be used in the preparation of the chimeric proteins and  
 CC their heterodimer complexes. The heterodimer complexes, which are useful  
 CC for testing potential promoters and inhibitors of the binding of  
 CC integrins to their ligands, function as blood platelet substitutes and  
 CC hemostatics and as diagnostic agents.  
 XX  
 SQ Sequence 963 AA;

Query Match 55.1%; Score 1253; DB 19; Length 963;  
 Best Local Similarity 51.6%; Pred. No. 3.9e-70;  
 Matches 273; Conservative 32; Mismatches 94; Indels 130; Gaps 15;  
 QY 1 ENFPKYLHYDETS--HOLLCD-----KCPGCT-----YLKQHC 33  
 DB 441 dsfkiplgftveevilygiceceqgsepshkchegntfecgacrcnrgvgrhc 500  
 QY 34 TAKMTVCAPCPDHY-----YTDWHTSDCLYCSVPKELQVVKQ-----EC 76  
 DB 501 ecstdevnsedmdayckrnsseicscgvcgvcvckrdntneisgkfgcecdnfc 560  
 QY 77 NRTN-----RVCEKRGYLETEFC--LKRSCTPP-----GEGVQAGTPER 117  
 DB 561 drsnglicggngvckervcecpnpytgscdcsldtstceasngqicnrgicecy 616  
 QY 118 NTVKRCPCDGFSTNETSSKAP-----CRKH-----TNCVSFGLLT 153  
 DB 617 --vck-ctdpxkfggqtcemcqtclgvcachkvcvqrafnkgektdtqecsyfnktv 673  
 QY 154 QKGNATHDNI-----CSGNSE-----STQKCGI-----D 177  
 DB 674 esrdklipqvpqdpvshckekdvddcwfyftysvngnvmvvhvenpecptgedpeep 733  
 QY 178 VTVDKTHTCPCPAPPELLCGPSVFLPPPKPKDTLMISRPETVCVVVDVSHEDPEVKFNW 237  
 DB 734 kscdkthtccpccpapellgppsvflfpkpkdtlmisrtpetvcvvvdvshedpevkfnw 793  
 QY 238 YVDGVEVHNAKTPREQYNSTRVYVSLTVLHODWLNCKEYKCKVSNKALPAPIEKTIS 297  
 DB 794 yvdgvevhnaktkpreeqynstnyrvsvltvclhqdwlngkeyckvckvsnkalpapieltis 853  
 QY 298 KAKGQPREQVYTPPSPRDLTNRQVSLTCLVKGFPSPDTAVESWNGQPPENNYKTTTPV 357  
 DB 854 kagqpreqvylppsrdeitknqysltclvkgfypsdlavewesngqppennnykttppv 913  
 QY 358 LDSGSGFFLYSKUTVDKSRWQQGNVPSGVSMHREALNHYTKSLSPG 406  
 DB 914 ldsdgsfflyskitvdksrwqqgnvpsgvsmhrealnhytqkslsisp 962

RESULT 13  
 W73513  
 ID W73513 standard; Protein; 388 AA.  
 XX  
 AC W73513;  
 DT 02-MAR-1999 (first entry)  
 XX  
 DE Rabbit TGFbetaRII:Fc protein.  
 XX  
 KW Transforming growth factor-beta receptor; TGF-beta receptor; arthritis;  
 KW fusion protein; fibroproliferative disorder; diabetic nephropathy;  
 KW glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;  
 KW collagen vascular disorder; therapy; rabbit.  
 XX  
 OS Oryctolagus sp.  
 XX  
 PN W09848024-A1.  
 XX  
 PD 29-OCT-1998.  
 XX  
 PF 16-APR-1998; 98WO-US07587.  
 XX  
 PR 18-APR-1997; 97US-0044641.  
 XX  
 PA (BIOI) BIOGEN INC.  
 XX  
 PI Cate R, Gotwals P, Koteliarsky V, Sanicola-Nadel M;  
 XX  
 DR WPI: 1998-609994/51.  
 DR N-PSDB; V08998.  
 XX  
 PT Transforming growth factor-beta receptor fusion protein - used to  
 PT treat fibroproliferative disorders  
 XX  
 PS Claim 4; Page 18-19; 70pp; English.  
 XX  
 CC This sequence is a fusion protein of the rabbit transforming growth  
 CC factor-beta receptor II (TGF-betaRII) and an antibody Fc region. The  
 CC encoded protein is an example of a protein of the invention, which are  
 CC isolated TGF-beta receptor fusion proteins that competitively inhibit  
 CC binding of TGF-beta to TGF-beta receptor. The fusion protein can be used  
 CC in a method for lowering the levels of TGF-beta in an individual having  
 CC arthritis. It can also be used to treat medical conditions such as  
 CC fibroproliferative disorders. The fibroproliferative disorder is a  
 CC kidney, intraocular or pulmonary fibrosis, especially selected from  
 CC diabetic nephropathy, glomerulonephritis, proliferative  
 CC vitreoretinopathy, or myelofibrosis. The fusion proteins can also be used  
 CC to treat collagen vascular disorders, including systemic sclerosis,  
 CC polymyositis, scleroderma, dermatomyositis, and systemic lupus  
 CC erythematosus. They can also be administered following coronary  
 CC angioplasty, to prevent restenosis or scarring and reclosing of arteries.  
 XX  
 SQ Sequence 388 AA;

Query Match 55.0%; Score 1250.5; DB 19; Length 388;  
 Best Local Similarity 65.2%; Pred. No. 2e-70;  
 Matches 257; Conservative 18; Mismatches 72; Indels 47; Gaps 10;  
 QY 37 WKTVCAPCPDHY-----TDSWHTSDE-----CLYC---SPVCKELQVVKQECNRTN 81  
 DB 17 wtiaistipphvdkvsnndmmvtdngavkfpqlckfdvrsstcdnqkscmncslt-- 74  
 QY 82 RVCEKRGYLETEFCLEKHSRCPGFGVQAGTPERTVCKRCPPD-----GFFSNETSS 135  
 DB 75 sice-----kahevqv---avwrkndenitietvchdtklayhgfledsas 118  
 QY 136 KAPCRKHTNCSVEFG---LLLTQKGNATHDNCISGNSESTQKCIDVTVDKTHCTPCPCAP 192  
 DB 119 pkcimkek--kvfgctfmcscstcdndhlfseevttsspd1---vdkthctpcpcap 173

QY 193 ELLGGPSVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVHNAKTKPR 252  
 Db 174 eLLgppsVflfppkpdltmIsrtpevtcvvvdshedpevkfnwYdgvvHNAKtkpr 233  
 QY 253 EEQYNSTRYVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTL 312  
 Db 234 eeqYnstrYrvsvltvlhQdWlNgkEyckvsnkAlpApIEktIskAkGgprEqvYtlp 293  
 QY 313 PSRDLTKNOVSLCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLT 372  
 Db 294 psrdeltknqVslclvkgfypsdiavewesngqpennYkTtpPvldsdgsfflyskltv 353  
 QY 373 DKSRWQGNVFCSCVMHEALHNHYTKLSLSPG 406  
 Db 354 dksrwqgnvfscsvmhealhnhytklsislspg 387

RESULT 14  
 Y54063  
 ID Y54063 standard; Protein; 388 AA.  
 XX AC Y54063;  
 XX DT 27-MAR-2000 (first entry)  
 XX DE Amino acid sequence of TGF-beta type II receptor variant/IgG1 fusion.  
 XX KW Splice variant; rabbit; transforming growth factor-beta; TGF-beta;  
 KW type II receptor; Fc portion; human IgG1; fusion protein; arthritis;  
 KW fibroproliferative disease; renal; intra-ocular; pulmonary; fibrosis;  
 KW diabetic nephropathy; glomerulonephritis; collagen vascular disease;  
 KW proliferative vitreoretinopathy; myelofibrosis; systemic sclerosis;  
 KW polymyositis; scleroderma; dermatomyositis; systemic lupus erythematosus;  
 KW restenosis; wound; connective tissue production; adhesion; scarring;  
 KW post-radiation fibrosis.  
 XX OS Synthetic.  
 OS Oryctolagus cuniculus.  
 OS Homo sapiens.  
 XX PN W09965948-AL.  
 XX PD 23-DEC-1999.  
 XX PF 16-JUN-1999; 99WO-US13629.  
 XX PR 16-JUN-1998; 98US-0089452.  
 XX PA (BIOJ ) BIOGEN INC.  
 XX PI Kotelliansky V, Gotwals P, Cate R, Sanicola-Nadel M;  
 XX DR WPI; 2000-106083/09.  
 DR N-PSDB; 245251, 245252.  
 XX New fusion protein of a splice variant of transforming growth  
 PT factor-beta receptor, for inhibiting the growth factor, e.g. in  
 PT treatment of fibrosis  
 XX PS Disclosure; Page 62-63; 69pp; English.  
 XX CC The present sequence represents a splice variant of a rabbit transforming  
 CC growth factor-beta (TGF-beta) type II receptor fused to the Fc portion  
 CC of human IgG1. The fusion protein has higher affinity for TGF-beta than  
 CC fusion proteins comprising the non-variant form of the receptor. The  
 CC fusion protein contains soluble TGF-beta receptor constructs that are  
 CC devoid of a transmembrane region (and are secreted from the cell) but  
 CC retain the ability to bind TGF-beta. The protein competitively inhibits  
 CC binding of TGF-beta to cellular receptors and/or forms an inactive  
 CC complex with TGF-beta. The protein is used to reduce levels of TGF-beta,  
 CC for treatment of arthritic conditions associated with overexpression  
 CC of TGF-beta, especially fibroproliferative diseases, e.g. renal,  
 CC intra-ocular or pulmonary fibrosis; diabetic nephropathy;

CC glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;  
 CC collagen vascular disease, e.g. systemic sclerosis, polymyositis,  
 CC scleroderma, dermatomyositis or systemic lupus erythematosus; and  
 CC fibrosis associated with restenosis. It is also used for treating  
 CC wounds, to prevent overproduction of connective tissue and so prevent  
 CC adhesions or scarring, and to prevent post-radiation fibrosis (by  
 CC administration to patients about to undergo radiation therapy).  
 XX SQ Sequence 388 AA;  
 Query Match 55.0%; Score 1250.5; DB 21; Length 388;  
 Best Local Similarity 65.2%; Pred. No. 2e-70;  
 Matches 257; Conservative 18; Mismatches 72; Indels 47; Gaps 10.  
 QY 37 WKTVCAPCPDHY-----TDSWHTSDE-----CLYC---SPVCKELQYVKQECNRTHN 81  
 Db 17 wtIastIpPhvqksvndmmvtdngavkfPqIckfcDvrsctcdnqscmsncsit-- 74  
 QY 82 RVCECKEGRYLEIEFLKLRSCPPGFGVWQAGTPERNTVCKRCPD-----GFFSNETSS 135  
 Db 75 sice-----kahevsv-----avrkndenitltvchdpklayhgfiledsas 118  
 QY 136 KAPCRKHTNCSVFG---LLLTQGNATHDNCISGNSSTQKCGIDVTVDKTHTCPPCPAP 192  
 Db 119 pKcimkek--kvfgetffmcsctdecdndhiifseeYttstspdl---vdkthtccppap 173  
 QY 193 ELLGGPSVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVHNAKTKPR 252  
 Db 174 eLLgppsVflfppkpdltmIsrtpevtcvvvdshedpevkfnwYdgvvHNAKtkpr 233  
 QY 253 EEQYNSTRYVSVLTVLHODWLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTL 312  
 Db 234 eeqYnstrYrvsvltvlhQdWlNgkEyckvsnkAlpApIEktIskAkGgprEqvYtlp 293  
 QY 313 PSRDLTKNOVSLCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLT 372  
 Db 294 psrdeltknqVslclvkgfypsdiavewesngqpennYkTtpPvldsdgsfflyskltv 353  
 QY 373 DKSRWQGNVFCSCVMHEALHNHYTKLSLSPG 406  
 Db 354 dksrwqgnvfscsvmhealhnhytklsislspg 387

RESULT 15  
 W70797  
 ID W70797 standard; protein; 592 AA.  
 XX AC W70797;  
 XX DT 03-FEB-1999 (first entry)  
 XX DE Human interleukin-6R-alpha-Fc.  
 KW gp130; cytokine antagonist; interleukin; gamma-interferon;  
 KW granulocyte macrophage colony-stimulating factor; J peptide;  
 KW transforming growth factor-beta.  
 XX OS Synthetic.  
 OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT Protein 1..358  
 FT /note= "human IL-R-alpha"  
 FT Misc-difference 2  
 FT /label= L2V  
 FT /note= "changed to accommodate a Kozak sequence"  
 FT Peptide 1..19  
 FT /note= "signal peptide"  
 FT Misc-difference 359..360  
 FT /note= "Ala-Gly bridge"  
 FT Protein 361..592  
 FT /note= "Fc domain of human IgG1"

FT Disulfide-bond 371..374

XX US5844099-A.

XX 01-DEC-1998.

XX 27-NOV-1995; 95US-0563105.

XX 27-NOV-1995; 95US-0563105.

PR 20-OCT-1993; 93US-0140222.

XX (REGE-) REGENERON PHARM INC.

XX Economides A, Stahl N, Yancopoulos GD;

XX WPI; 1999-044669/04.

XX Cytokine antagonists - comprising extracellular domains of

PT specificity-determining and signal-transducing components of

PT cytokine receptor

XX Example 3; Fig 5; 46pp; English.

XX The present sequence represents the amino acid sequence of human

CC interleukin (IL)-6 $\alpha$ -Fc. The protein is used in the course

CC of the invention. The specification describes cytokine antagonists

CC comprising only the extracellular domain of the specificity-determining

CC component of the cytokine receptor and the extracellular domain of a

CC signal-transducing component of the cytokine receptor. The cytokine

CC is an interleukin (IL-1, IL-2, IL-3, IL-4, IL-5 or IL-15),

CC granulocyte macrophage colony-stimulating factor (GM-CSF),

CC gamma-interferon or transforming growth factor-beta (TGF-beta). The

CC antagonist is capable of binding the cytokine to form a nonfunctional

CC complex. The compounds have therapeutic activity as cytokine antagonists

CC and can also be used in assays for identifying novel agonists and

CC antagonists of cytokines.

XX SQ Sequence 592 AA;

Query Match 55.0%; Score 1250; DB 20; Length 592;

Best Local Similarity 84.0%; Pred. No. 3.4e-70;

Matches 241; Conservative 10; Mismatches 28; Indels 8; Gaps 3;

QY 125 PDGFFSNETSSKAPCRKHTNCVFGLLLTOKGNATHDNIC---SGNSESTQKCGIDVT 179

Db 308 peamgtptwtesrppaenevstpmqalttnkdd---dnllfrdsanatsipvdagepks 364

QY 180 VDKTHTCPPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 239

Db 365 cdkthtccpccpapelggpsvflfpkpkdtlmisrtpevtcvvdvshedpevkfnwv 424

QY 240 DGEVHNNAKYPREEQNSYRYRVSVLTIVLHQDLNCKEYKCKVSNKALPAPIEKTISK 299

Db 425 dgevhnaktpreeqnystyrsvltivhqdwlngkyckvsnkalpapiektiska 484

QY 300 KGPQREPQVTLPSRDELTKNVSITCLVKGYFSPVDIAVWESNGQPENNYKTTPPVLD 359

Db 485 kgprepqvytlppsrdeitknqvsitclvkgyfspdviawesngqpennnyktppvld 544

QY 360 SDGSFFLYSLKLTVDKSRWQGNFVSCVMHEALHNHYTKQSLSLSPG 406

Db 545 sdgsfflyskltvdksrwqgnvfscvmhealhnhytkqslslspg 591

Search completed: March 1, 2001, 09:17:06

Job time: 224 sec





**SPTREMBL 15: \***

QY	1	ETPPKYLHYDEETSHQLLCKDKCPPGTYLKQHCTAKWKTCAPCPDHYHYTDSWHTSDECL	60
Db	1	ETPPKYLHYDEETSHQLLCKDKCPPGTYLKQHCTAKWKTCAPCPDHYHYTDSWHTSDECL	60
QY	61	YCSPVCKELQYVQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVYVQAGTIPERNIV	120
Db	61	YCSPVCKELQYVQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVYVQAGTIPERNIV	120
QY	121	CKRCPPGFFSNETSSKAPCRKHTNCSVFGLLLTQGNATHDNCISGNSBESTQKSG	175

Db 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFLGLLLTQGNATHDNCNCSGSESTQKCG 175

RESULT 2  
000300 PRELIMINARY; PRT; 401 AA.

AC 000300; 060236;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)  
GN (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER IIB).  
OS TNFRSF11B OR OPG OR OCIF.  
OC Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX TISSUE=KIDNEY;  
RX MEDLINE=97262071; PubMed=9108485;  
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,  
Ra Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,  
Ra Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,  
Ra Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,  
Ra Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,  
Ra Suggs S., Boyle W.J.;  
RT "Osteoprotegerin: a novel secreted protein involved in the regulation  
of bone density.";  
RL Cell 89:309-319(1997).  
RN [2]  
RC SEQUENCE FROM N.A.  
RX TISSUE=LUNG FIBROBLAST;  
RX MEDLINE=98151033; PubMed=9492069;  
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,  
Ra Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,  
Ra Tsuda E., Morinaga T., Higashio K.;  
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and  
osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits  
osteoclastogenesis in vitro.";  
RL Endocrinology 139:1329-1337(1998).  
RN [3]  
RC SEQUENCE FROM N.A.  
RX TISSUE=PLACENTA;  
RX MEDLINE=98351569; PubMed=9688283;  
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;  
RT "Cloning and characterization of the gene encoding human  
osteoprotegerin/osteoclastogenesis-inhibitory factor.";  
RL Eur. J. Biochem. 254:685-691(1998).  
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY  
SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY  
SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,  
KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN  
A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN  
THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.  
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
DR EMBL; AB002146; BAA25910.1; -;  
DR EMBL; AB008822; BAA32076.1; -;  
DR EMBL; AB008821; BAA32076.1; JOINED.  
DR EMBL; U94332; AAB53709.1; -;  
DR HSP; P25942; ICDF.  
DR MIM; 602643; -;  
DR INTERPRO; IPR001368; -;  
DR PFAM; PF00020; TNFR\_C6; 3.  
DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
DR PROSITE; PS00050; TNFR\_NGFR\_2; 2.  
DR PRODOM; PD000771; -; 1.

KW Glycoprotein; Repeat; Cytokine; Signal.  
FT SIGNAL 1 21  
FT CHAIN 22 401  
FT DOMAIN 23 183  
FT REPEAT 23 63  
FT REPEAT 64 106  
FT REPEAT 107 143  
FT REPEAT 144 201  
FT DOMAIN 306 365  
FT DISULFID 41 54  
FT DISULFID 44 62  
FT DISULFID 65 80  
FT DISULFID 83 97  
FT DISULFID 87 105  
FT DISULFID 118 142  
FT DISULFID 145 160  
FT CARBOHYD 98 98  
FT CARBOHYD 152 152  
FT CARBOHYD 165 165  
FT CARBOHYD 178 178  
FT CARBOHYD 289 289  
FT CONFLICT 263 263  
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 44.7%; Score 1013; DB 4; Length 401;  
Best Local Similarity 99.4%; Pred. No. 2.6e-81;  
Matches 174; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ETPPKYLYHDEETSHQLLCDKCPGTYLKQHCCTAKWKTVCAPCPDHYTDSWHTSDSCL 60  
Db 22 ETPPKYLYHDEETSHQLLCDKCPGTYLKQHCCTAKWKTVCAPCPDHYTDSWHTSDSCL 81

Qy 61 YCSPVKRELQVQECNTHNRVCECKEGRYLEIEFCLKHKRSCPPGFGVQAGTPERTV 120  
Db 82 YCSPVKRELQVQECNTHNRVCECKEGRYLEIEFCLKHKRSCPPGFGVQAGTPERTV 141

Qy 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFLGLLLTQGNATHDNCNCSGSESTQKSG 175  
Db 142 CKRCPDGFSSNETSSKAPCRKHTNCSVFLGLLLTQGNATHDNCNCSGSESTQKCG 196

RESULT 3  
008712 PRELIMINARY; PRT; 401 AA.

AC 008712; 070202;  
DT 01-JUL-1997 (TREMBlrel. 04, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)  
GN (OCIF).  
GN TNFRSF11B OR OPG.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX STRAIN=BALB/C; TISSUE=KIDNEY;  
RX MEDLINE=97262071; PubMed=9108485;  
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,  
Ra Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,  
Ra Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,  
Ra Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,  
Ra Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,  
Ra Suggs S., Boyle W.J.;  
RT "Osteoprotegerin: a novel secreted protein involved in the regulation  
of bone density.";  
RL Cell 89:309-319(1997).  
RN [2]  
RC SEQUENCE FROM N.A.  
RX STRAIN=129/OLA, AND NIH SWISS;  
RX MEDLINE=98382527; PubMed=9714833;



```
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
SQ SEQUENCE 401 AA; 46192 MW; FEC6A31F1D4E573A CRC64;

Query Match 38.8%; Score 879; DB 11; Length 401;
Best Local Similarity 85.1%; Pred. No. 1.7e-69;
Matches 149; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 ETTPPKYLHYDETHSHOLLCDKCPGTYLKHOCTAKWTKVCAPCPDHYHDTSDWHTSDECL 60
   ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 22 ETTPPKYLHYDETHSHOLLCDKCPGTYLKHOCTAKWTKVCAPCPDHYHDTSDWHTSDECV 81
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 61 YCSPVKCELVQKQECNTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVVOAGTPERTV 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 82 YCSPVKCELVQKQECNTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVVOAGTPERTV 141
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 121 CKRCPDGFNETSSKAPCRKHTNCVFGLLLTOKGNATHDNCISGSESTQKSG 175
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142 CKRCPDGFNETSSKAPCRKHTNCVFGLLLTOKGNATHDNCISGSESTQKSG 196
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
Q9RIA4 PRELIMINARY; PRT; 437 AA.
AC Q9RIA4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE GAMMAL HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SQ SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv)".
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -
DR HSP; P01842; 7FAB.
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; ig; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 36.2%; Score 819.5; DB 11; Length 437;
Best Local Similarity 62.8%; Pred. No. 3.2e-64;
Matches 140; Conservative 43; Mismatches 35; Indels 5; Gaps 2;

QY 184 CPP--CPAPELLGSPSVLEFPKPKDITMISRTPEVTCVVVDVSHEDPEVKFNMYDQVE 241
   ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 KCPCICITPEV---SSVFIFPPPKVDLTITLPKVTCTVVDVSKDDPEVFQSFHVDVE 273
   ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 242 VHNAKTPREQYNSTYRVSVLTVLHODWLNGKEYCKYKSNKALPAPIETKISKAKGP 301
   ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 274 VHTAQTPREQYNSTYRVSVLTVLHODWLNGKEYCKYKSNKALPAPIETKISKAKGP 333
   ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 302 REPOVYTLPSRDELTKNOYSLCLVKGFFPSDITAVESNGQENNYKTPPVLDSDGS 361
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 334 KAPOVYTIPTPKQEMAKDKVSLTCHMITFFPEDITVWQNGQPAENYKNTQPIMDTDS 393
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 362 FFYLSKLTVDKSRWQGNVFSVCVNMHEALHNHYTKQSLSPG 404
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 394 YFVYSKLVNQKSWNEACNTTCSVLHESLHNHHTKSLSPG 436
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
```

```

RESULT 7
Q9PUSO ID Q9PUSO PRELIMINARY; PRT; 302 AA.
AC Q9PUSO
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE DCOV RECEPTOR.
OS Salvelinus fontinalis (Brook trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN [1]
RP SEQUENCE FROM N.A.
RA Bobe J., Goetz F.W.;
RT "A tumor necrosis factor receptor homolog is up-regulated in the brook
trout (Salvelinus fontinalis) ovary at the completion of ovulation.";
RL Biol. Reprod. 0:0-0(1999).
DR EMBL; AF156738; AAD56428.1; -.
DR HSSP; P19438; 1EXT.
DR INTERPRO; IPR000561; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE. 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match 18.0%; Score 406.5; DB 13; Length 302;
Best Local Similarity 43.1%; Pred. No. 5e-28;
Matches 72; Conservative 32; Mismatches 60; Indels 3; Gaps 2;

QY 11 DEETSHOLLCDKCPPGTYLKQHTAKWTKVACPCPDHYTDSWHTSDECLYSPVKRELQ 70
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 27 DSGLSIVDCRCPPGTYLRAPCSAMRKSDCAECPCNGAYTEFNHHSKLCRCS-MCAENQ 85
Y 71 YVQECNTHNVCECKEGRYLE--EIFELKHSRCPGPGVQAGTPERNVTCKRCPDGF 128
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 86 VVQECSPSNCECEKEGYFNNKACIKHKECPCPGYANTTGTGPHQDTECVQCOAGF 145
Y 129 FSNETSKAPCRKHTNCSVFGLLLTKGNATHDNICSGNSESTQKSG 175
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 146 YSEVSSAKATCLAQSNCKVGLRVLLKQDWHNTLCASYDLKTRDG 192

RESULT 8
Q16042 ID Q16042 PRELIMINARY; PRT; 439 AA.
AC Q16042
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-JAN-1999 (TReMBLrel. 09, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91370690; PubMed=1966549;
RA Dembic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R.,
RA Brockhaus M., Lesslauer W.;
RT "Two human TNF receptors have similar extracellular, but distinct
intracellular, domain sequences.";
RL Cytokine 2:231-237(1990).
DR EMBL; S63368; AAB19824.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.

```

```

DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
DR PRODOM; PD000771; -.
SQ SEQUENCE 439 AA; 46090 MW; FEBCEB329CC67FE6 CRC64;

Query Match 14.9%; Score 337.5; DB 4; Length 439;
Best Local Similarity 26.6%; Pred. No. 9.4e-22;
Matches 107; Conservative 44; Mismatches 163; Indels 89; Gaps 15;

QY 9 HYDETSHOLLCDKCPPGTYLKQHTAKWTKVACPCPDHYTDSWHTSDECLYSPVKCE 68
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 23 YIDO--TAQMCCSKCSPQHAQVCTKTSDTVCDCSDSTYTQLMNVWPECLSCGSRCS 80
Y 69 LOYVVKQECNTHNVCECKEGRYLEI-----EFCLKHSRCPGPGVQAGTPERNVTCK 122
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 81 DQVETQACTREQNRICTCRPGWYCALSKQECRCICAPLKRCPGFGVARPGTETSDVVCK 140
Y 123 RCPDGFESNETSSKAPCRKHTNCSVFGLLLTKGNATHDNICSGNSESTQKSGG----- 177
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 141 PCAPGTFSTNTSSDIDCRPHQICNVAI-----PGNASMDAVCTSTSPTRSNAPGAVHLPQ 196
Y 178 -GGGGTCTPPCPAPE-----LLGGPSVFLPPPKPKDT-----LMISRT----- 215
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 197 PVSTRSQHTOPTPEPTAPSTFSLPMGPS-----PPAEGSTGDFALPVGLIVGVTALGLL 252
Y 216 ---EVTCCVVVDVSHEDP-----EVKFNWYVDGVEVHNNAK-TKPREQYNSTYRVVSVLTV 266
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 253 IIGVNCVIMTQVKKKPLCLQREAK---VPHLPADKARGTGGPEQQH-----LLIT 300
Y 267 LHODWLNKGEYKCKVSNKALPAPIETKISKAKGPREPOVYVTPPSRDELTK----- 318
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 301 APSSSSSLSASSALDRAP-----TRNQPPQPGVEASGAGERASTGSSDSSPG 351
Y 319 ---NQVSLTCLVKGFPDIAVEWESN-----GQPENNYKTP 353
D : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 352 GHGTQVNVTCIVNVCSSSHSSQCSQASSTMGTDTSPPSPS 394

RESULT 9
Q62327 ID Q62327 PRELIMINARY; PRT; 459 AA.
AC Q62327
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=95178848; PubMed=7873884;
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
gene.";
RL Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -.
DR HSSP; P19438; 1NCF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
FT NON_TER 1
FT VARIANT 87 87 S -> T.
FT VARIANT 93 93 T -> I.
FT VARIANT 268 268 F -> I.
FT VARIANT 345 345 S -> F.
FT VARIANT 421 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

```

Query Match	14.8%	Score	335;	DB	11;	Length	459;
Best Local Similarity	24.6%	Pred.	No. 1.6e-21;				
Matches	100;	Conservative	51;	Mismatches	160;	Indels	96;
Gaps	12;						
QY	9	HYDEETSHQLCDKCPGTYLKOHCTAKWKTVCAPCPDHYVYTDSWHTSDECLYCSPVCKE	68				
DB	31	YDRKA--QMCCKACPPGQVVKVFKCNKTSDTVCADCBASWYQVWNOFRTCLSCSSCGST	88				
QY	69	LQYVQECNTRHNRVCECKEGRYLEIF-----CLKHRSCPPGFGVQVQAGTPERNTVC	121				
DB	89	DOVETRACTRQONRVCAACEAGRYCALKTHSGCRQCCKMLSKGPGFGVGSASSRPNGNVLC	148				
QY	122	KRCPDGFSSNETSKAPCRKHTNCVSFGLLTQKGNATHNICSGNEES-----	170				
DB	149	KACAPGTFTDSTSDVCRPHRICSLAI-----PGNASTDAVCAPESPTLSAIPRTLTVS	204				
QY	171	-----TQKSGGGGGGGTCCPPCAPPELL-----GGPSV-----	198				
DB	205	QPEPTRSQPLDQEPG-----PSQTSILTSLGSTPIIEQSTKGGISLPIGLIVG	254				
QY	199	-----FLFPPPKDFTLMSRPEVTCVVVVDVSHEDPEVKFNWYDGVVEVHNAKTKPREBQY	254				
DB	255	TSGLMLGLVNCVFIWQRKXKPSCLORDA--KVPHYPDEKSKQDAVGL-----EQQH	304				
QY	255	NSTYRVVSVLTVLHQDWLNCKEYCKVSNKALPA---PIEKTSIKAKGPREPQVYTLPP	311				
DB	305	-----LTTAPSSSSSSLESASACDRRAPPGCHPQARVMAEAOGSQBARASSRISD	356				
QY	312	SRDELTKNOVSLTCLVKGYFSPDIAVEWSN-----QGPENNYKTPP	353				
DB	357	SSHGSHGTHVNVTCIVNCVSSSDHSSOCSSOASATVGDPPDAKPSASP	403				

```

RESULT 10
O88734 PRELIMINARY; . PERT; 482 AA.
ID O88734;
AC O88734;
DT DT O1-NOV-1998 (T-EMBLrel. 08, Created)
DT DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
DE P80 TNF-ALPHA RECEPTOR.
GN TNFR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hurler B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;
RT "The Mouse Tumor Necrosis Factor Receptor Gene:Genomic Structure and
RT Characterization of the two Transcripts.";
RL Genomics 0:0-0(0).
DR EMBL; Y14619; CAA74969.1; -.
DR EMBL; Y14620; CAA74969.1; JOINED.
DR EMBL; Y14621; CAA74969.1; JOINED.
DR EMBL; Y14622; CAA74969.1; JOINED.
DR EMBL; Y14623; CAA74969.1; JOINED.
DR EMBL; Y14679; CAA74969.1; JOINED.
DR HSSP; PJ9438; INCF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS0050; TNFR_NGFR_2; 3.
DR PRODOM; PD000771; -.
DR PRODOM; PD000771; -.
SQ SEQUENCE 482 AA; 51106 MW; F6C15046B48FF83C CRC64;

```

Query Match	14.5%;	Score 328;	DB 11;	Length 482;
Best Local Similarity	24.5%;	Pred. No. 7.2e-21;		
Matches 102;	Conservative 48;	Mismatches 160;	Indels 106;	Gaps 13;

QY 9 HYDEETSHQLLCKPCPPGTYLKQHCT-----AKWKTVCAPCPDHYTDSWHTSDECLY 61

Db	46	YDRKA	-OMQCAKCPGGYVHKFNKTS	SDTVCADSDTVCADCEASMYQWNQFTCLS	103
QY	62	CSPVCKELQYVQECNRTNRNCECKEGRYLEIEF	-----CLKHRSPPGFGVGVQAGT	114	
Db	104	CSSSCSTDQVETRAC	TQKNRYCACACAGRYCALKTHSGSCROCMRLSCGPGFGVASSRA	163	
QY	115	PERNTVCKRCPDGFP	SNETHSSKAPCKKHTNCVSFGLLLTKQGNATHDNTCSGNSES	170	
Db	164	PNGNVLKACAFGT	SSDTSSDTCVDRPHRICSLAI	219	
QY	171	-----	TKSGGGGGGGTCDPCPAPELL	198	
Db	220	PRTLVSQPEPTRSQPLDQPG	-----PSQTPSILTSLGSTPIIEQSTKGISLP	269	
QY	199	-----	FLFPKPKDTLMISTPEVTCVVVD	345	
Db	270	IGLIVGVTSLGLMLGLVNC	FILVQRKKPSCILORDAKVPHV	28	
QY	246	KTKPREEQYNSTRYVS	VLTLHQDWLNGKEYCKVSNKALPA	302	
Db	329	TTAPSSSSSL	-----ESSASGDRRAPGGHQPQARVHAEAGSOE	370	
QY	303	EPQVYTLPPSRDELTKNOV	SLTKLVKGYFSPDI	353	
Db	371	ARASSRISDSSHGSGHGT	HNVTICVNCSSDHS	426	

[illegible]

RN	[2]	SEQUENCE FROM N.A.	
RP		Parker A.;	
RA		Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.	
RL		EMBL; AF068868; AAC34583.1; -.	
RR		EMBL; AL096801; CAB75692.1; -.	
DR		HSP: P07174; INGR.	
DR		INTERPRO: IPR000488; -.	
DR		INTERPRO; IPR001368; -.	
DR		PFAM; PF00020; TNFR_c6; 4.	
DR		PFAM; PF00531; death; 1.	
DR		PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.	
DR		PROSITE; PS50017; DEATH_DOMAIN; 1.	
DR		PROSITE; PS50050; TNFR_NGFR_2; 1.	
DR		PRODOM; PD000771; -; 1.	
DR		Receptor.	
SW		655 AA; 71844 MW; 48939391CA4852A33 CRC64;	
KO		SEQUENCE	

Query Match	13.2%;	Score 298;	DB 4;	Length 655;
Best Local Similarity	27.8%;	Pred. No. 4.6e-18;		
Matches	87;	Conservative	44;	Mismatches 132;
		Indels	50;	Gaps 11;

QY 7 YLHYDEETSHQLCDKCPPGTYLKQHCTAKWKTVCAPCPDHYYTDSWHTSDECLYCPVC 66

```
Db 54 YRVDNRATGVLCDKCPAGTYVSEHCTNLSRVCSGPCVGTETREHENGIEKCHDCSQPC 113
QY 67 KELQYVQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVVOAGTPTERNVTKRCPCD 126
Db 114 PWFMIKELPCALTDRECTCPGWFQSNATCAPTVCPVGWGVKKGKTETEDVRCQCAR 173
QY 127 GFSTNETSSKAPCKRHTNCSVFLGILLTKGNATHDNICSGNSESTOKSGGGGGGTCTP- 185
Db 174 GTFSDVPSSVMCKAVTDCLSQNLVVIKPGTKETDNVC-----GTLPS 216
QY 186 -----PCPAPELLGSPVLEFPKPKDTLMSLRTPETVCVVVDVSHED-----PDVKF 233
Db 217 FSSSTSPSP-----GTALF-----PRP-EHMETHEVPSSTYVPKGNSTNESASVRKV-L 267
QY 234 NKWVDGVEVHNATKPREQYNSTYRVSVTLVLDHDLNGKEVKCKVSNKALPAPI--- 290
Db 268 SSTQEGTVPNTSSARGKEDVNT--LPNLQVNVHQ---QGPHRHL--KLLPSMEATG 320
QY 291 -EKTISKAKQPR 302
Db 321 GEKSTPIKPKR 333
```

```
RESULT 12
Q9NPP6 PRELIMINARY; PRT; 416 AA.
AC Q9NPP6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RT "The European IMAGE Consortium for Integrated Molecular analysis of
RT human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389978; CAB97534.1; -
FT NON_TER 1
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;
```

```
Query Match 11.9%; Score 269.5; DB 4; Length 416;
Best Local Similarity 25.8%; Pred. No. 8.6e-16;
Matches 90; Conservative 42; Mismatches 136; Indels 81; Gaps 13;

QY 104 PPGFGVVOAGTPTERNVTKRC- PDGFFSNETSSKAPCKHTNCSVFGLLLTQKGNATHDN 162
Db 82 PKVFPILSDTPDQGNVVAACLVQGFPPQP-----LSVTWSESGQN 123
QY 163 ICSGNSESTQKSGG-----GGGGGTC-----PPCPAPPELL 193
Db 124 VTARNPPSDASGDLTYTTSQLTLPATQCPDGKSVTCHVKHTNPSQDVTVPVCPV--- 180
QY 194 GGSVFLFPKP-----KDTLMSLRTPETVCVVVDVSHEDPEVKFNWVDGVE 241
Db 181 -----PPPCCHPLSLHRLPALEDLLGSEANLTCTLTGL-RDASGATFTVPSGK 231
QY 242 VHNATKPREQYNSTYRVSVTLVLDHDLNGKEVKCKVSNKALPAPIEKTISKAKGP 301
Db 232 --SAVGPPERDLICGYSVSVLPGCAQPNWGHETFTCTAAHKLTPLTANITKS--GNT 288
QY 302 REPQVYTLPPSRDELTKNQ-VSLTCLVKGFYPSDIAEVWESNGQ--PENNYKTTTPPVLD- 357
```

```
Db 289 FRPEVHLLPPPSSEALNELVTLTCLARGFSPKDVLRWLGQSQELPREKYLTLTWASQEP 348
QY 358 SDG--SFFLYSKLTVDKSRWQGVNFSCSVNHEALHNHYTKSLSPG 404
Db 349 SQGTTTFAVTSILRVAADWKKGDTFCMVGHEALPLAFTQKTIDRLAG 397

RESULT 13
Q90544 PRELIMINARY; PRT; 684 AA.
AC Q90544;
DT 01-NOV-1996 (TREMBlrel. 01; Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OX NCBI_TaxID=7801;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RX MEDLINE=95183140; PubMed=7877689;
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,
RA Flajnik M.F.;
RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks.";
RL Nature 374:168-173(1995).
DR EMBL; U18701; AAB48195.1; -
DR HSSP; P01857; 1FC1.
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; Ig; 6.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Signal.
FT SIGNAL.
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DEF CRC64;
```

```
Query Match 11.9%; Score 269.5; DB 13; Length 684;
Best Local Similarity 26.2%; Pred. No. 1.6e-15;
Matches 100; Conservative 53; Mismatches 162; Indels 67; Gaps 16;

QY 51 DSWHTSDECLYCS-----PVCKELOYVQECNRTHNRVCECKEGRYLEIEFCLKRS 102
Db 315 EEWQSGVE-YTCSAKQDQSTPVYKTRKARVEPTKPHRL----- 354
QY 103 CPFGVVOAGTPTERNVTKRC PDGFFSNETS---SKAPCKHTNCSVFGLLLTQKGNAT 159
Db 355 LPSPPEIQSTSATLTCLIR---GFYPDKVSVWQKDDVSVSANVTNFTALQ--DLT 409
QY 160 HDNICSGNSESTQ-KSGGGGGGTCTCP-----CPAPELGGSPVFLPPP 203
Db 410 FSTRSLNLTAWEKSGAKYCTCTASHPPSQSTVKRVIRNOKVDCROTDI---SVSLKP 465
QY 204 KPQDTLMSLRTPETVCVVVDVSHEDPE-VKFNWYVDGVEVHNATKPREQYNSTYRVVS 262
Db 466 -PFEIWTQQTATVCEIV---YSDLENIKVFWQVNGVERKKGVETQNPWMSGSKSTIVS 521
QY 263 VLVVLDHDLNGKEVKCKVSNKALPAPIEKTISKAK-GQPREPOVYTLPPSRDEL-TKNQ 320
Db 522 KLVKMASEWDSGTEYVCLVEDSELTPTVKASIRKANVYQMHPPKVIYLLHPTSDIDTENS 581
QY 321 VSLTCLVKGFYPSDIAEVWESNGQ-PENNYKTTTPPVLDSDGSEFFLYSKLAVDKSRWQGN 379
Db 582 ATLMCLATNPHPAEIVYGVGMANDTLDDSGYRTQVDSEKSGSSFTVDRRLRTAAEWSDT 641
QY 380 VFCSSVMHEALHN---HYTKS 398
Db 642 TYSCLVGHPSLNRDLIRSTNKS 663
```

```
RESULT 14
O9UP60 PRELIMINARY; PRT; 384 AA.
AC O9UP60;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE SNC73 PROTEIN.
GN SNC73.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zheng S., Cao J., Cao W., Cai X., Geng L.;
RA "Identification and characterization of SNC73, a gene which is down-
RA regulated in colorectal cancer.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067420; AAC19365.1;
DR HSSP; P01825; 7FAB.
DR INTERPRO; IPR003006;
DR PFAM; PF00047; 19; 3.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 384 AA; 40947 MW; BA7ADC3CA5A9DD48 CRC64;

Query Match 11.9%; Score 268.5; DB 4; Length 384;
Best Local Similarity 26.2%; Pred. No. 9.6e-16;
Matches 89; Conservative 40; Mismatches 138; Indels 73; Gaps 12;

Qy 115 PERNTVCKRCPDGFPSNETSSKAPCRKHTNCVFLGLLTQKGNATHDNCSESTQKS 174
Db 49 PDGNVVIACLVQGFPPQEP-----LSVTWSESGQGVGTARNFPPSODA 90
Qy 175 GGG-----GGGGTFC-----PPCAPPELLGGPSVFLFPKP 205
Db 91 SGDLTSSQLTLPATOCGLAGKSVTCHVKYHTNPQSDVTVCVPSTPTPTSPST-PTTP 149
Qy 206 K-----DYLMSRPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTKPR 250
Db 150 SPSCCHPRLSLHRPALELLGSEANLCTLTGL-RDASGVTFWTTPSSGK--SAVQGGP 206
Qy 251 EBYQNSTYRVVSVTLVHODWLNQKEYCKVSKNALPAPIEKTISKAKGQPREPOVYTL 310
Db 207 ERLDCCGYSVSVLPGCCAEPNHHGKTFTCTAAYPESKTPLTATLSKSGNTFRPEVHL 265
Qy 311 PSRDELTKNQ-VSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTTTPPVLD-SDG--SPFL 364
Db 266 PSEELALNELVTLTCLARGFSPKDLVRLQGSQGLPREKYLTLWASRQEPSQGTTF 325
Qy 365 YSKLTVDKRWQGNVFCVSWHEALHHNYTKSLSPG 404
Db 326 TSILRVAEDMKGGDTFCWVGHEALPLAFTQKTDRLAG 365

RESULT 15
O35305 PRELIMINARY; PRT; 625 AA.
AC O35305;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE RECEPTOR ACTIVATOR OF NF-KAPPA-B PRECURSOR (TNF-RELATED ACTIVATION-
DE INDUCED CYTOKINE RECEPTOR) (RANK).
GN TNFRSF11A OR RANK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC TISSUE=LIVER EPITHELIUM;
RX MEDLINE=98032977; PubMed=9367155;
RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
RA Tometsko M.E., Roux E.R., Teepe M.C., DuRose R.F., Cosman D.,
RA Galibert L.;
RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
RT and dendritic-cell function.";
RL Nature 390:175-179(1997).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS RANKL.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: CONTAINS A LA-MNGPR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; AF019046; AAB86810.1;
DR HSSP; P25942; 1CDF.
DR MGD; MGI:1314891; Tnftrsf11a.
DR INTERPRO; IPR000561;
DR INTERPRO; IPR001368;
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 1.
DR PRODOM; PD000771; 1.
KW Receptor; Glycoprotein; Transmembrane; Signal; Repeat.
FT SIGNAL 1 30
FT CHAIN 31 625
FT DOMAIN 31 214
FT TRANSMEM 215 235
FT DOMAIN 236 625
FT DOMAIN 34 196
FT REPEAT 34 70
FT REPEAT 71 114
FT REPEAT 115 153
FT REPEAT 154 196
FT DISULFID 35 47
FT DISULFID 48 61
FT DISULFID 51 69
FT DISULFID 72 87
FT DISULFID 93 113
FT DISULFID 115 125
FT DISULFID 127 134
FT DISULFID 128 152
FT DISULFID 155 170
FT DISULFID 176 195
FT CARBOHYD 106 106
FT CARBOHYD 175 175
SQ SEQUENCE 625 AA; 66621 MW; F8C1872E99511D8E CRC64;

Query Match 11.2%; Score 253.5; DB 11; Length 625;
Best Local Similarity 25.3%; Pred. No. 3.6e-14;
Matches 103; Conservative 41; Mismatches 146; Indels 117; Gaps 18;

Qy 20 CDKCPPTGLYKQHTAKWKTVCAPCDHYVTDSWHTSDECLYCSPVC---KELQVVKQEC 76
Db 48 CSRCEPGKYLSSKCTPTSDSVCLPCGDEYLDTWNEEDKCL-LHKVCDAGKALVAV-DPG 105
Qy 77 NRTNRCVCEKGRY--LEIEFLCKHRSCPGFGVQVQAGTPERNTVCKRCPDGFESNETS 134
Db 106 NHTAPRCACACTAGYHWNDSCECCRRNTECAPGFGAQAHPLOLNKDVCTPCLLGFSDVFS 165
Qy 135 SKAPCRKHTNCSVFGLLLTQKGNATHDNCSESTQKSGGGGGGGTCTPPCPAPPELLG 194
Db 166 STDCKPWTNCTLLGKLEAHQGTSTSDVVCSS-SMTLRR-----PPKEAQAYL- 212
Qy 195 GPSVFLFPKPKDILMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTKPREQY 254
Db 213 -PSLIVL-----LLFISVVVVAIIIFG-----YVRKGGK----- 241
Qy 255 NSTVRVSVLTVLHODWLN-----GKEY---KCKVSNKA-----LPADIE 291
Db 242 -----ALTANLWNVNDACSSLGKNKSSGDRCSGSHSATSSQOEVCIGILLMTREE 293
Qy 292 KTIK-----AKQGP-----REPOVYTL-----PSRDELTKNQVTS 322
```



Db 294 KMVPEDGAGVCGPVCAAGGPAEVRDSRTFTLVSEVETQGDLSRKIPTEDXYTDRFSQPS 353  
Qy 323 LTCIV-----KGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFF 363  
Db 354 TGSLLLIQOGSKSIPFPQEPLEVGENDSLQCFGTGTSTVDSEGCDF 400

Search completed: March 1, 2001, 09:20:11  
Job time: 409 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:47 ; Search time 40.97 Seconds  
(without alignments)  
318.448 Million cell updates/sec

Title: US-09-389-782A-7  
Perfect score: 2264  
Sequence: 1 ETEPPKYLHYDEETSHQLC.....VMHEALHNHYTQKSLSPG 404

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	1204	53.2	330	1	GC1_HUMAN	P01857	homo sapien
2	1137.5	50.2	326	1	GC2_HUMAN	P01859	homo sapien
3	1133	50.0	327	1	GC4_HUMAN	P01861	homo sapien
4	1125.5	49.7	290	1	GC3_HUMAN	P01860	homo sapien
5	909	40.2	323	1	GC_RABIT	P01870	oryctolagus
6	890.5	39.3	329	1	GC2_CAVPO	P01862	cavia porce
7	841.5	37.2	329	1	GC3_MOUSE	P22436	mus musculus
8	840	37.1	333	1	GC6_RAT	P20761	rattus norv
9	835.5	36.9	398	1	GC3M_MOUSE	P03987	mus musculus
10	824.5	36.4	326	1	GC1_RAT	P20759	rattus norv
11	812.5	35.9	324	1	GC1_MOUSE	P01868	mus musculus
12	812.5	35.9	393	1	GC1M_MOUSE	P01869	mus musculus
13	804.5	35.5	329	1	GC3_MOUSE	P20762	rattus norv
14	804	35.5	330	1	GCA_MOUSE	P01863	mus musculus
15	804	35.5	399	1	GCAM_MOUSE	P01865	mus musculus
16	797	35.2	335	1	GCAB_MOUSE	P01864	mus musculus
17	780.5	34.5	322	1	GCA_RAT	P20760	rattus norv
18	774	34.2	336	1	GC6_MOUSE	P01866	mus musculus
19	774	34.2	405	1	GC6_MOUSE	P20762	rattus norv
20	360.5	15.9	391	1	MUCB_HUMAN	P04220	homo sapien
21	353.5	15.6	421	1	EPC_MOUSE	P06336	mus musculus
22	353	15.6	429	1	EPC_RAT	P01855	rattus norv
23	353	15.6	454	1	MUC_HUMAN	P01871	homo sapien
24	352.5	15.6	455	1	MUC_MOUSE	P01872	mus musculus
25	347.5	15.3	476	1	MUCM_MOUSE	P01873	mus musculus
26	343	15.2	428	1	EPC_HUMAN	P01854	homo sapien
27	339.5	15.0	461	1	TNR2_HUMAN	P20333	homo sapien
28	337	14.9	458	1	MUC_RABIT	P03988	oryctolagus
29	332	14.7	479	1	MUCM_RABIT	P04221	oryctolagus
30	331.5	14.6	454	1	MUC_MESAU	P06337	mesocricetu
31	330.5	14.6	457	1	MUC_SUNMU	P20768	suncus mur
32	330.5	14.6	474	1	TNR2_MOUSE	P25119	mus musculus
33	329.5	14.6	450	1	MUC_CANFA	P01874	canis famil

## ALIGNMENTS

RESULT 1	GC1_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG GAMMA-1 CHAIN C REGION.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RL	Biochemistry 9:3161-3170(1970).			
RN	[3]			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RL	Biochemistry 9:3171-3181(1970).			
RN	[4]			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The chymotryptic peptides of the H-chain, alignment of the tryptic peptides and discussion of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RN	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[6]			
RP	DISULFIDE BONDS			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X. Intrachain disulfide bonds."			

34	304.5	13.4	438	1	HVC2_HETFR	P23085	heterodontu
35	298	13.2	438	1	HVC3_HETFR	P23087	heterodontu
36	295	13.0	299	1	ALC_RABIT	P01879	oryctolagus
37	294	13.0	461	1	HVCM_HETFR	P23088	heterodontu
38	287	12.7	393	1	HVC3_HETFR	P23086	heterodontu
39	284	12.5	289	1	CD40_MOUSE	P27512	mus musculus
40	282.5	12.5	446	1	MUC_CHICK	P01875	gallus gall
41	277.5	12.3	370	1	HVC1_HETFR	P23084	heterodontu
42	269.5	11.9	340	1	ALC2_HUMAN	P01877	homo sapien
43	268.5	11.9	353	1	ALC1_HUMAN	P01876	homo sapien
44	267	11.8	353	1	ALC1_GORGO	P20758	gorilla gor
45	264	11.7	277	1	CD40_HUMAN	P25942	homo sapien

Biochemistry 9:3188-3196(1970).  
 [7]  
 RN DISULFIDE BONDS.  
 RX MEDLINE=77070267; PubMed=1002129;  
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 RT "Rule of antibody structure. The primary structure of a monoclonal  
 RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and  
 RT characterization of the protein, the L- and H-chains, the  
 RT cyanogen bromide cleavage products, and the disulfide bridges.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
 [8]  
 RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RX MEDLINE=81208100; PubMed=7236608;  
 RA Deisenhofer J.;  
 RT "Crystallographic refinement and atomic models of a human Fc fragment  
 RT and its complex with fragment B of protein A from Staphylococcus  
 RT aureus at 2.9- and 2.8-A resolution.";  
 RL Biochemistry 20:2361-2370(1981).  
 CC GIM(1) MARKERS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE  
 CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)  
 CC MARKER & THE GIM (NON-1) MARKERS.  
 CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC 35,116,198,269 & 272.  
 CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
 CC 268-272.  
 CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC RESIDUES 198,267&272.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/annouce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J00228; AAC82527.1; ALT\_INIT.  
 DR PIR; A02146; GHU.  
 DR PDB; 1FC1; 15-JUL-92.  
 DR PDB; 1FC2; 15-JUL-92.  
 DR MIM; 147100;  
 DR INTERPRO; IPR000495;  
 DR INTERPRO; IPR003006;  
 DR PFAM; PF00047; Ig; 3.  
 DR PROSITE; PS00290; IG\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT DOMAIN 99 110 HINGE.  
 FT DOMAIN 111 223 CH2.  
 FT DOMAIN 224 330 CH3.  
 FT DISULFID 27 83  
 FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT CARBOHYD 180 180  
 FT VARIANT 97 97  
 FT VARIANT 239 239  
 FT VARIANT 241 241  
 FT MOD\_RES 330 330  
 FT STRAND 123 126  
 FT HELIX 130 134  
 FT TURN 136 137  
 FT STRAND 141 148  
 FT STRAND 158 162  
 FT TURN 163 164

FT STRAND 165 166  
 FT STRAND 175 178  
 FT STRAND 183 190  
 FT HELIX 193 197  
 FT TURN 198 199  
 FT STRAND 202 206  
 FT STRAND 215 219  
 FT STRAND 227 227  
 FT STRAND 230 234  
 FT HELIX 238 240  
 FT TURN 241 242  
 FT STRAND 245 256  
 FT STRAND 260 266  
 FT TURN 267 268  
 FT STRAND 269 270  
 FT STRAND 274 276  
 FT STRAND 280 281  
 FT TURN 283 284  
 FT STRAND 287 296  
 FT HELIX 297 301  
 FT TURN 302 303  
 FT STRAND 306 312  
 FT TURN 313 314  
 FT TURN 316 317  
 FT STRAND 320 324  
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
 Query Match 53.2%; Score 1204; DB 1; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 6.3e-74;  
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 183 TCPPCPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 242  
 Db 108 TCPPCPAPELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 167  
 QY 243 HNAKTKPREQYNSTYRVVSVLTVLHQDLNGLKEYCKVSNKALPAPKEKTISKAKGQPR 302  
 Db 168 HNAKTKPREQYNSTYRVVSVLTVLHQDLNGLKEYCKVSNKALPAPKEKTISKAKGQPR 227  
 QY 303 EPQVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEESNGQPENNYKTPPVLDSDGSF 362  
 Db 228 EPQVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEESNGQPENNYKTPPVLDSDGSF 287  
 QY 363 FLYSKLTVDKSRWQQGNVSCVMHEALHNYHTOKSLSPG 404  
 Db 288 FLYSKLTVDKSRWQQGNVSCVMHEALHNYHTOKSLSPG 329  
 RESULT 2  
 GC2\_HUMAN  
 ID GC2\_HUMAN STANDARD; PRT; 326 AA.  
 AC P01859;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-2 CHAIN C REGION.  
 GN IGHG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82197621; PubMed=6804948;  
 RA Ellison J.W., Hood L.E.;  
 RT "Linkage and sequence homology of two human immunoglobulin gamma  
 RT heavy chain constant region genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).  
 RN [2]  
 RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
 RX MEDLINE=81007873; PubMed=6774012;  
 RA Wang A.-C., Tung E., Fudenberg H.H.;  
 RT "The primary structure of a human IgG2 heavy chain: genetic,



FT DOMAIN 221 327 CH3.  
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 83  
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 141 201  
FT DISULFID 247 305  
SQ SEQUENCE 327 AA; 35940 MW; 38DBD811EF208E7A CRC64;

Query Match 50.08; Score 1133; DB 1; Length 327;  
Best Local Similarity 92.98; Pred. No. 3.5e-69;  
Matches 208; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 181 GGTCCPAPELGSPVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKENYVDGV 240  
Db 103 GPPCSCPAPELGSPVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKENYVDGV 162  
QY 241 EVHNAKTPREEQNSTYRVSVLTVLHODWLNGKEYCKVSNKALPAPIETISKAKGQ 300  
Db 163 EVHNAKTPREEQNSTYRVSVLTVLHODWLNGKEYCKVSNKALPAPIETISKAKGQ 222  
QY 301 PREQVYTLPSRDELTKNQVSLCLVKGFPYSDIAVWESNGOPENNYKTPPVLDSDG 360  
Db 223 PREQVYTLPSRDELTKNQVSLCLVKGFPYSDIAVWESNGOPENNYKTPPVLDSDG 282  
QY 361 SFFLYSKLTVDKSRWQGNVFCSCVMHEALNHYTKQSLSLSPG 404  
Db 283 SFFLYSKLTVDKSRWQGNVFCSCVMHEALNHYTKQSLSLSPG 326

## RESULT 4

GC3\_HUMAN  
ID GC3\_HUMAN STANDARD; PRT; 290 AA.  
AC P01860;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).  
GN IGHG3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
[1]  
RA SEQUENCE (DISEASE PROTEIN WIS).  
RA MEDLINE-81021548; PubMed-6774747;  
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;  
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:  
RT gamma 3 heavy-chain disease protein Wis.";  
RL Biochemistry 19:4304-4308(1980).  
RN [2]  
RP REVISIONS TO 12-97 OF PROTEIN WIS.  
RX MEDLINE-77118561; PubMed-402363;  
RA Michaelson T.E., Frangione B., Franklin E.C.;  
RT "Primary structure of the 'hinge' region of human IgG3. Probable  
RT quadruplication of a 15-amino acid residue basic unit";  
RL J. Biol. Chem. 252:883-889(1977).  
RN [3]  
RP REVISIONS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).  
RX MEDLINE-77021516; PubMed-823945;  
RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;  
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.  
RT Structure of the FC fragment of immunoglobulin G3";  
RL Biochem. Biophys. Res. Commun. 71:907-914(1976).  
RN [4]  
RP SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).  
RX MEDLINE-82247835; PubMed-6808505;  
RA Alexander A., Steinmetz M., Barriault D., Frangione B.,  
RA Franklin E.C., Hood L., Buxbaum J.N.;  
RT "Gamma heavy chain disease in man: cDNA sequence supports partial  
RT gene deletion model";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).  
CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA

CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11  
CC NORMALLY PRESENT IN THE HINGE REGION.  
CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.  
CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE  
CC REF.2.  
CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION  
CC AND ALL OF THE CH1 REGION.  
CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL  
CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL  
CC GAMMA-3 HEAVY CHAINS.  
CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM  
CC OR ANOTHER GAMMA CHAIN SUBCLASS.  
CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR  
CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE  
CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE  
CC SEGMENT (12-28).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: J00231; AAA52805.1; ALT\_SEQ.  
CC PIR: A02149; G3HUI.  
CC MIM: 147120;  
CC INTERPRO: IPR000495; -;  
CC INTERPRO: IPR003006; -;  
CC PFAM: PF00047; ig; 2.  
CC PROSITE: PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
FT DOMAIN 12 73  
FT HINGE.  
FT DOMAIN 74 183  
FT DISULFID 24 24 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 27 27 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 33 33 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 39 39 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 42 42 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 48 48 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 54 54 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 57 57 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 63 63 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 69 69 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT DISULFID 72 72 INTERCHAIN (WITH HEAVY CHAIN DIMER).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC...).  
FT MOD\_RES 290 290 REMOVED POST-TRANSLATIONALLY.  
FT VARIANT 126 127 QV -> EB (IN ZUC).  
FT VARIANT 134 134 /FTid=VAR\_003890.  
FT VARIANT 139 134 P -> L (IN OMM).  
FT VARIANT 139 139 /FTid=VAR\_003891.  
FT VARIANT 182 182 F -> Y (IN OMM).  
FT VARIANT 182 182 /FTid=VAR\_003892.  
FT VARIANT 227 227 T -> A (IN OMM).  
FT VARIANT 227 227 /FTid=VAR\_003893.  
FT VARIANT 227 227 S -> N (IN OMM).  
FT VARIANT 227 227 /FTid=VAR\_003894.  
FT VARIANT 279 279 MISSING (IN ZUC).  
FT VARIANT 279 279 /FTid=VAR\_003895.  
FT SEQUENCE 290 AA; 32331 MW; 569BC95705B2F46 CRC64;  
SQ

Query Match 49.7%; Score 1125.5; DB 1; Length 290;

Best Local Similarity 69.6%; Pred. No. 9.8e-69;  
Matches 217; Conservative 19; Mismatches 23; Indels 53; Gaps 5;

QY 97 CLKHRSCT---PPGFGVQAGTPERNITCKPCPDGFSNETSSKAP-CRKHTNCSVFGLLL 152  
Db 27 CPKPKSCDTPPP-----CPKCPBP-KSCDTPPPCPKCPKPKSCDT----- 65

QY 153 TQKGNATHDNICSGNSESTQKSGGGGGGTGCPCPAPPELLGGPSVFLPPKPKDTLMIS 212  
Db 66 -----PPCPCPCAPPELLGGPSVFLPPKPKDTLMIS 97

QY 213 RTEVTCVVVDVSHEDPEVKFNKVDGVEVHNKATKPREQYNSTYRVSVLTVLHODWL 272  
Db 98 RTEVTCVVVDVSHEDPEVQFKYVVDGVQVHNKATKPREQYFNSTFRVSVLTVLHODWL 157

QY 273 NGREYCKVSNKALPAPIETISKAKGPREPOVYTLPPSRDELTKNOVSLTCLVKGFP 332  
Db 158 DGRYCKVSNKALPAPIETISKAKGPREPOVYTLPPSRDELTKNOVSLTCLVKGFP 217

QY 333 SDIAVESNGOPENNYKTPPPVLDSDGSPFLYSLKTLVDKSRWQOGNPFSCSVMEALHN 392  
Db 218 SDIAVESNGOPENNYKTPPPVLDSDGSPFLYSLKTLVDKSRWQOGNPFSCSVMEALHN 277

QY 393 HYTKLSLSPG 404  
Db 278 RFTQKLSLSPG 289

RESULT 5

GC\_RABBIT STANDARD; PRT; 323 AA.

ID GC\_RABBIT  
AC P01870;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DE IG GAMMA CHAIN C REGION.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84030930; PubMed=6313520;  
RA Bernstein K.E., Alexander C.B., Mage R.G.;  
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant  
F-1 haplotype."  
RL Immunogenetics 18:387-397(1983).  
RN [2]  
RP SEQUENCE OF 1-128.  
RX MEDLINE=76135469; PubMed=1243651;  
RA Pratt D.M., Mole L.E.;  
RT "Sequence studies on the constant region of the Fd sections of rabbit  
immunoglobulin G of different allotype."  
RL Biochem. J. 151:337-349(1975).  
RN [3]  
RP SEQUENCE OF 88-266 FROM N.A.  
RX MEDLINE=83299917; PubMed=6193512;  
RA Martens C.L., Moore K.W., Steimetz M., Hood L., Knight K.L.;  
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma  
heavy chain and identification of two genomic C gamma genes."  
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).  
RN [4]  
RP SEQUENCE OF 132-161.  
RX MEDLINE=70110015; PubMed=5461106;  
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;  
RT "Sequence studies of the Fd section of the heavy chain of rabbit  
immunoglobulin G."  
RL Biochem. J. 116:249-259(1970).  
RN [5]  
RP SQUILL OF 129-131 AND 155-322.  
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;  
RL (in) Killefer J. (eds.);  
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,  
Stockholm (1967).

CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,  
CC 104-THR, AND REF.5 THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15  
CC MARKERS AND REF.5 THE E15 MARKER.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
CC -----  
CC EMBL: M16426; AAA31289.1; -  
DR PIR: A02161; GHRB.  
DR INTERPRO: IPR000495; -  
DR INTERPRO: IPR003006; -  
DR PFAM: PF00047; Ig; 3.  
DR PROSITE: PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT VARIANT 104 104 T -> M (IN D11 MARKER).  
FT VARIANT 185 185 T -> A (IN E15 MARKER).  
FT CONFLICT 48 48 N -> E (IN REF. 2).  
FT CONFLICT 71 71 V -> VPV (IN REF. 2).  
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).  
FT CONFLICT 173 173 N -> D (IN REF. 5).  
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).  
FT CONFLICT 201 201 N -> D (IN REF. 5).  
FT CONFLICT 218 218 Q -> E (IN REF. 5).  
FT CONFLICT 233 233 E -> Q (IN REF. 5).  
FT CONFLICT 246 246 N -> D (IN REF. 5).  
FT CONFLICT 256 256 E -> G (IN REF. 5).  
FT CONFLICT 260 260 N -> D (IN REF. 5).  
FT CONFLICT 266 266 E -> D (IN REF. 5).  
FT CONFLICT 280 280 Y -> W (IN REF. 5).  
FT CONFLICT 284 284 N -> S (IN REF. 5).  
SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;

Query Match 40.2%; Score 909; DB 1; Length 323;  
Best Local Similarity 72.8%; Pred. No. 3.3e-54;  
Matches 163; Conservative 28; Mismatches 31; Indels 2; Gaps 1;

QY 183 TC--PPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENYVDGV 240  
Db 99 TCSKPTCPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENYVDGV 158

QY 241 EVHNAKTPREEQYNSTYRVSVLTVLHODWLNGKCKYCKVSNKALPAPIETISKARGQ 300  
Db 159 QVTRAPPLREQQPNSTIRVSVLTPLTHQDWLRCGKCKVHNKALPAPIETISKARGQ 218

QY 301 PREQVTLPPSRDELTKNOVSLTCLVKGFPDIAVESNGOPENNYKTPPPVLDSDG 360  
Db 219 PLEPKVYTMGPPEELSLRSVSLTCMNGFYPDSDISVEWEKNGKAEDNYKTPPAVLSDG 278

QY 361 SFELYSLKTLVDKSRWQOGNPFSCSVMEALHNHYTKLSLSPG 404  
Db 279 SYFLYNKLSVPTSEWQGRDVFSCVMEALHNHYTKLSLSPG 322

RESULT 6

GC2\_CAVPO STANDARD; PRT; 329 AA.

ID GC2\_CAVPO  
AC P01862;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2 CHAIN C REGION  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
RN [1]  
RP SEQUENCE OF 1-3.









DB 104 CKPCICTVPEV---SSVFIPPPKDKVLTITLTPKVCVVVDISKDDPEVQFSWFVDVDE 160  
 QY 242 VHNAKTKPREEOYNSTYRVVSVLTVLHODWLNGKEYKCKVSNKALPAPIEKTISKAGQP 301  
 DB 161 VHTAQTPREOQNSFRSVSLPIHQDWLNGKEFKCRVNSAFAPIEKTISKTKGRP 220  
 QY 302 REPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGQPNENYKTTTPVLDSDGS 361  
 DB 221 KAPOVVTIPPPKQMAKDKVSLTCMTDFPEDITVWQNGQPAENYKNTQPIMTNGS 280  
 QY 362 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPG 404  
 DB 281 YFYSLKLVQKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPG 323

## RESULT 12

GC1M\_MOUSE STANDARD; PRT; 393 AA.  
 AC P01869;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=80045036; PubMed=115593;  
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
 RA Takahashi N., Mano Y.;  
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin  
 gamma 1 chain gene";  
 RL Cell 18:559-568(1979).  
 RN [2]  
 RP SEQUENCE OF 323-393 FROM N.A.  
 RX MEDLINE=82197626; PubMed=6804950;  
 RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;  
 RT "mRNA for surface immunoglobulin gamma chains encodes a highly  
 conserved transmembrane sequence and a 28-residue intracellular  
 domain";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).  
 RN [3]  
 RP SEQUENCE OF 323-366 FROM N.A.  
 RX MEDLINE=82115295; PubMed=6799207;  
 RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,  
 RA Eisenberg D., Wall R.;  
 RT "Gene segments encoding transmembrane carboxyl termini of  
 immunoglobulin gamma chains";  
 RL Cell 26:19-27(1981).  
 RN [4]  
 RP SEQUENCE OF 1-44 FROM N.A.  
 RX MEDLINE=8222190; PubMed=6283537;  
 RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;  
 RT "Nucleotide sequences of gene segments encoding membrane domains of  
 immunoglobulin gamma chains";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).  
 CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA  
 SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED  
 GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-  
 BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED  
 IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND  
 SEGMENT OF MU CHAINS.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 -----

DR EMBL; V00793; CAA24172.1; -  
 DR EMBL; V00793; CAA24173.1; -  
 DR EMBL; V00793; CAA24174.1; -  
 DR EIR; B02153; GLMSW.  
 DR MGD; MGI:96446; IGH-4.  
 DR INTERPRO: IPR000495; -  
 DR INTERPRO: IPR003006; -  
 DR PFAM: PF00047; I9; 3.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW Alternative splicing; Transmembrane.  
 FT NON\_TER 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 110 HINGE.  
 FT DOMAIN 111 217 CH2.  
 FT DOMAIN 218 324 CH3.  
 FT DISULFID 27 82  
 FT DISULFID 102 102  
 FT DISULFID 104 104  
 FT DISULFID 107 107  
 FT DISULFID 109 109  
 FT DISULFID 138 138  
 FT CARBOHYD 174 174  
 FT DISULFID 244 302  
 FT TRANSMEM 340 357  
 FT DOMAIN 358 393  
 FT SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;  
 QY 184 CPP--CPAPELLGGPSVFLPPPKDPTLMISRTPETVTVVVDVSHEDPEVKFNWYDVGVE 241  
 DB 104 CKPCICTVPEV---SSVFIPPPKDKVLTITLTPKVCVVVDISKDDPEVQFSWFVDVDE 160  
 QY 242 VHNAKTKPREEOYNSTYRVVSVLTVLHODWLNGKEYKCKVSNKALPAPIEKTISKAGQP 301  
 DB 161 VHTAQTPREOQNSFRSVSLPIHQDWLNGKEFKCRVNSAFAPIEKTISKTKGRP 220  
 QY 302 REPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGQPNENYKTTTPVLDSDGS 361  
 DB 221 KAPOVVTIPPPKQMAKDKVSLTCMTDFPEDITVWQNGQPAENYKNTQPIMTNGS 280  
 QY 362 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPG 404  
 DB 281 YFYSLKLVQKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPG 323

## Query Match

Best Local Similarity 62.3%; Score 812.5; DB 1; Length 393;  
 Matches 139; Conservative 44; Mismatches 35; Indels 5; Gaps - 2;

## RESULT 13

GCC\_RAT STANDARD; PRT; 329 AA.  
 AC P20762;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-2C CHAIN C REGION.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88166903; PubMed=3127222;  
 RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;  
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant  
 region cDNA: extensive homology to mouse gamma 3";  
 RL Eur. J. Immunol. 18:317-319(1988).  
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way



ID GCAM\_MOUSE STANDARD; PRT; 399 AA.  
AC P01865;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2A CHAIN C REGION, MEMBRANE-BOUND FORM.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-8222190; PubMed-6283537;  
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;  
RT "Nucleotide sequences of gene segments encoding membrane domains of  
immunoglobulin gamma chains";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).  
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA  
SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED  
GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-  
BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED  
IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND  
SEGMENT OF MU CHAINS.  
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE  
IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF  
THE A ALLELE.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J00471; AAB59661.1; ALT\_INIT.  
DR PIR; A02154; G2MSAM.  
DR MGI; 96443; IGH-1.  
DR INTERPRO: IPR000495; -  
DR INTERPRO: IPR003006; -  
DR PFAM; PF00047; ig; 3.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
KW Transmembrane; Alternative splicing.  
FT NON\_TER 1 15 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 15 15  
FT DISULFID 27 82  
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 144 204  
FT DISULFID 250 308  
FT DISULFID 346 363 POTENTIAL.  
FT TRANSMEM 346 399 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 364 399 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 180 180  
SQ SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FE0 CRC64;

Query Match 35.5%; Score 804; DB 1; Length 399;  
Best Local Similarity 65.0%; Pred. No. 4.4e-47;  
Matches 145; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

QY 184 CPP--CPAPELGGPSVFLPPPKPTLMISRTPEVTCVVVDYSHEDPEVKFNKYVDGVE 241  
DB 107 CPPCKCAPNLLGGPSVFIIPPKIKDVLMSLSPITVCVVVDVSEDDPDVQISWVNNVE 166  
QY 242 VHNAKTKPREEQNYSYRVSVLTVLHODWLNGLNGKEYCKVSKNKPALPAPIEKTISKAKGP 301  
DB 167 VHTAQITQTHREDYNSLTVRVVSALPIQHDWMSGKEFKCKVNNKDLPAPIERTISKPKGSV 226  
QY 302 REPQVYTLPPSRDELTKNQVSLTCLVKGYFSPDAIIEWESNQGPENNYKTTTPVLDSDGS 361  
DB 227 RAPQVYVLPPEEMTKKQVTLTCMVTDMPEDIIYVWNTNNGKTELNYKNTEPVLDSGS 286

QY 362 FFYLSKLTVDKSRWQGNVFSQVMHEALHNHYTKQSLSPG 404  
DB 287 YFMYSKLRVEKKNNWVERNSYSCSVVHEGLNHHHTTKSFRTPG 329

Search completed: March 1, 2001, 09:17:48  
Job time: 266 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:48 ; Search time 79.26 Seconds  
(without alignments)  
346.100 Million cell updates/sec

Title: US-09-389-782A-7  
Perfect score: 2264  
Sequence: 1 ETFPPKYLHYDEETSHQLLC.....VMHEALHNHYTKQSLSLSPG 404

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues  
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1204	53.2	255	4 S31866	Ig gamma-1 chain C
2	1204	53.2	330	1 GHU	Ig heavy chain V r
3	1202	53.1	374	2 S72664	Ig heavy chain V r
4	1198	52.9	374	2 S69339	Ig gamma chain C r
5	1162	51.3	234	2 PT0207	Ig gamma-3 chain C
6	1146.5	50.6	377	2 A23511	Ig gamma-3 chain C
7	1144.5	50.6	377	2 A60764	Ig gamma-2 chain C
8	1137.5	50.2	326	1 G2HU	Ig gamma-4 chain C
9	1133	50.0	327	1 G4HU	Ig gamma-3 heavy c
10	1125.5	49.7	289	1 G3HUW1	Ig gamma chain C r
11	909	40.2	323	1 GHRB	Ig gamma 2b chain
12	902.5	39.9	328	2 I47160	Ig gamma 2a chain
13	902.5	39.9	328	2 I47159	Ig gamma 4 chain c
14	900.5	39.8	277	2 I47162	Ig gamma-2 chain c
15	890.5	39.3	329	1 G2GP	Ig gamma-2 chain c
16	880.5	38.9	328	2 I47158	Ig gamma 1 chain c
17	874.5	38.6	328	2 I47161	Ig gamma 3 chain c
18	850.5	37.6	470	2 S22080	Ig heavy chain pre
19	841.5	37.2	329	1 G3MSC	Ig gamma-3 chain C
20	841	37.1	308	2 C30554	Ig heavy chain C r
21	841	37.1	472	2 S31459	Ig gamma-1 chain -
22	840	37.1	333	2 PS0018	Ig gamma-2b chain
23	835.5	36.9	398	1 G3MSM	Ig gamma-3 chain C
24	824.5	36.4	326	2 PS0017	Ig gamma-1 chain C
25	822.5	36.3	444	2 PC4136	monoclonal antibody
26	812.5	35.9	324	1 G1MS	Ig gamma-1 chain C
27	812.5	35.9	393	1 G1MSM	Ig gamma-1 chain C
28	804.5	35.5	329	2 S00847	Ig gamma-2c chain
29	804	35.5	330	1 G2MSA	Ig gamma-2a chain

30	804	35.5	399	1 G2MSAM	Ig gamma-2a chain
31	804	35.5	469	2 S37483	Ig gamma-2a chain
32	797	35.2	335	1 G2MSAB	Ig gamma-2a chain
33	794	35.1	446	2 S40295	Ig gamma-2a chain
34	780.5	34.5	322	2 PS0019	Ig gamma-2a chain
35	774	34.2	336	1 G2MS11	Ig gamma-2b chain
36	774	34.2	405	1 G2MSBM	Ig gamma-2b chain
37	774	34.2	474	2 S25057	Ig gamma-2b chain
38	760.5	33.6	327	2 S06611	Ig gamma-2 chain C
39	752	33.2	475	2 S01321	Ig gamma-2b chain
40	702	31.0	180	2 I46732	Ig gamma heavy cha
41	577.5	25.5	249	2 S69340	Ig heavy chain VHI
42	574.5	25.4	218	2 A36040	Ig heavy chain V-I
43	566	25.0	152	2 S14236	Ig gamma-1 chain C
44	502	17.8	572	2 B46529	Ig Y heavy chain (
45	361	15.9	549	2 S04845	Ig Y heavy chain pre

ALIGNMENTS

RESULT 1

S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli  
C:Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000  
C:Accession: S31866

R:Filpula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene produc

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FIL>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 53.2%; Score 1204; DB 4; Length 255;  
Best Local Similarity 100.0%; Pred. No. 3e-68;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 TCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 242

Db 33 TCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 92

QY 243 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPTEKTSKAKGQPR 30

Db 93 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPTEKTSKAKGQPR 15

QY 303 EPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEVSNQGPENNYKTTTPVLDSDGSF 362

Db 153 EPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEVSNQGPENNYKTTTPVLDSDGSF 212

QY 363 FLYSKLTVDKSRWQOGNWFSCVMHEALHNHYTKQSLSPG 404

Db 213 FLYSKLTVDKSRWQOGNWFSCVMHEALHNHYTKQSLSPG 254

RESULT 2

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A30564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma gene.

A:Reference number: A93433; MUID:82274238

A:Accession: A93433

A:Molecule type: DNA  
A:Residues: 1-330 <ELL>  
A:Cross-references: EMBL:Z17370  
A:Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers,  
R:Harris, L.J.  
submitted to the EMBL Data Library, October 1992  
A:Reference number: S33904  
A:Accession: S36861  
A:Molecule type: DNA  
A:Residues: 2-330 <HAR>  
A:Cross-references: EMBL:Z17370  
R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.  
Cell 29, 671-679, 1982  
A:Title: Structure of human immunoglobulin gamma genes: Implications for evolution of a  
A:Reference number: S33887; MUID:83001943  
A:Accession: S33887  
A:Molecule type: DNA  
A:Residues: 88-113:235-330 <TAK>  
A:Cross-references: EMBL:Z17370  
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,  
Biochemistry 9, 3161-3170, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen  
A:Reference number: A90563; MUID:71064024  
A:Contents: myeloma protein Eu  
A:Accession: B90563  
A:Molecule type: protein  
A:Residues: 1-96,'R',98-135 <CUN>  
A:Note: This sequence has the G1m(3) marker, 97-Arg  
R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
Biochemistry 9, 3171-3181, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequen  
A:Reference number: A90564; MUID:71064025  
A:Contents: Eu  
A:Accession: A90564  
A:Molecule type: protein  
A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',240,  
A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met  
R:Ponstingl, H.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976  
A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nle),  
igen Primärstruktur.  
A:Reference number: A91668; MUID:77070269  
A:Contents: myeloma protein Nie  
A:Accession: B91668  
A:Molecule type: protein  
A:Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27  
A:Note: this sequence has the G1m(17) and G1m(1) markers  
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOI  
A:Reference number: A91723; MUID:83289131  
A:Contents: myeloma protein KOI; disulfide bonds  
A:Accession: A91723  
A:Molecule type: protein  
A:Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH  
A:Note: this sequence has the G1m(3) and G1m(non-1) markers  
R:Gall, W.E.; Edelman, G.M.  
Biochemistry 9, 3188-3196, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid  
A:Reference number: A90565; MUID:71064027  
A:Contents: annotation; disulfide bonds  
R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976  
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob  
enbromide cleavage products, and the disulfide bridges.  
A:Reference number: A91667; MUID:77070267  
A:Contents: annotation; disulfide bonds  
C:Genetics:  
A:Gene: GDB:IGH1  
A:Cross-references: GDB:120085; OMTM:147100  
A:Map position: 14q32.33-14q32.33  
A:Introns: 99/1; 114/1; 224/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (h  
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM1>  
F:137-206/Domain: immunoglobulin homology <IM2>  
F:243-310/Domain: immunoglobulin homology <IM3>  
F:27-83,144-204,250-308/Disulfide bonds: #status experimental  
F:103/Disulfide bonds: interchain (to light chain) #status experimental  
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 53.2%; Score 1204; DB 1; Length 330;  
Best Local Similarity 100.0%; Pred. No. 4e-68;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 TCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 242  
Db 108 TCPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEV 167

QY 243 HNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPR 302  
Db 168 HNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPR 227

QY 303 EPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVWESNGOPENNYKTTTPVLDSDGSF 362  
Db 228 EPQVYTLPPSRDELTKNOVSLTCLVKGFPSDIAVWESNGOPENNYKTTTPVLDSDGSF 287

QY 363 FLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPG 404  
Db 288 FLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPG 329

RESULT 3  
S72664  
Ig heavy chain V region precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-2000  
C:Accession: S72664  
R:Khamilchi, A.A.  
submitted to the EMBL Data Library, September 1994  
A:Reference number: S72664  
A:Accession: S72664  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <KHA>  
A:Cross-references: EMBL:X81695  
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 53.18%; Score 1202; DB 2; Length 374;  
Best Local Similarity 90.8%; Pred. No. 6e-68;  
Matches 226; Conservative 5; Mismatches 10; Indels 8; Gaps 2;

QY 156 GNATHDNICNSBSTOKSGGGGGGCTPCPAPELLGGPSVFLFPPPKDRLMISRT 215  
Db 133 GGGTLVTVCS-EPKCDKT-----HTCPCPAPELLGGPSVFLFPPPKDRLMISRT 184

QY 216 EVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGK 275  
Db 185 EVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGK 244

QY 276 EYCKVSNKALPAPIEKTISKAKGQPRQVYTLPPSRDELTKNOVSLTCLVKGFPSDI 335  
Db 245 EYCKVSNKALPAPIEKTISKAKGQPRQVYTLPPSRDELTKNOVSLTCLVKGFPSDI 304

QY 336 AVEWESNGOPENNYKTTTPVLDSDGSFELYSKLTVDKSRWQGNVFCSSVMHEALHNHYT 395  
Db 305 AVEWESNGOPENNYKTTTPVLDSDGSFELYSKLTVDKSRWQGNVFCSSVMHEALHNHYT 364

QY 396 QKSLSLSPG 404  
|||||



Db 365 QKSLSLSPG 373

RESULT 4

Ig heavy chain V region precursor - human

C:Species: Homo sapiens (man)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 17-Mar-2000

C:Accession: S69339

R:Khamlichi, A.A.; Accouturier, P.; Preud'homme, J.L.; Cogne, M.

Eur. J. Biochem. 229, 54-60, 1995

A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.

A:Reference number: S69339; MUID:95262687

A:Accession: S69339

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <KHA>

A:Cross-references: EMBL:X81695

C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 52.9%; Score 1198; DB 2; Length 374;  
Best Local Similarity 99.1%; Pred. No. 1.1e-67;  
Matches 220; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 183 TCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEV 242

Db 152 TCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEV 211

QY 243 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 302

Db 212 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 271

QY 303 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSF 362

Db 272 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSF 331

QY 363 FLYSKLTVDKSRWQGNVFNCSVMHEALHNNHYTQKSLSPG 404

Db 332 FLYSKLTVDKSRWQGNVFNCSVMHEALHNNHYTQKSLSPG 373

RESULT 5

PT0207

Ig gamma chain C region - chimpanzee

C:Species: Pan troglodytes (chimpanzee)

C:Date: 23-Nov-1991 #sequence\_revision 23-Nov-1991 #text\_change 16-Jul-1999

C:Accession: PT0207

R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.

Mol. Immunol. 28, 319-322, 1991

A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.

A:Reference number: PT0207; MUID:91287716

A:Accession: PT0207

A:Molecule type: mRNA

A:Residues: 1-234 <EHR>

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 51.3%; Score 1162; DB 2; Length 234;  
Best Local Similarity 99.1%; Pred. No. 1.1e-65;  
Matches 214; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 183 TCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEV 242

Db 19 TCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEV 78

QY 243 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 302

Db 79 HNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR 138

QY 303 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSF 362

Db 139 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSF 198

QY 363 FLYSKLTVDKSRWQGNVFNCSVMHEALHNNHYTQKS 398

Db 199 FLYSKLTVDKSRWQGNVFNCSVMHEALHNNHYTQKS 234

RESULT 6

A23511

Ig gamma-3 chain C region (allotype G3m(b)) - human

C:Species: Homo sapiens (man)

C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 23-Jul-1999

C:Accession: A23511

R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.

Nucleic Acids Res. 14, 1779-1789, 1986

A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene.

A:Reference number: A23511; MUID:86148507

A:Accession: A23511

A:Molecule type: DNA

A:Residues: 1-377 <HUC>

A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056

C:Genetics:

A:Gene: GDB:IGHG3

A:Cross-references: GDB:119339; OMIM:147120

A:Map position: 14q32.33-14q32.33

A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: immunoglobulin

F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 50.6%; Score 1146.5; DB 2; Length 377;  
Best Local Similarity 66.0%; Pred. No. 1.7e-64;  
Matches 227; Conservative 19; Mismatches 33; Indels 65; Gaps 7;

QY 76 CNRTHNRCVCEKGYLEIEF-----CLKHRSC---PPFGVVOAGTPEPNTVC 121

Db 83 CNYNH-KPSNTKVDKRVELKTPGLDTHTCRCPPEKSCDTPP-----C 126

QY 122 KRCPDGFSSNETSKAP-CRKHTNCVFGLLLTOKGNATHNICNSGSESTQKSGGGGG 180

Db 127 PRCPPEP-KSCDTPPCPCPCPEKSCDT----- 152

QY 181 GGTCCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGV 240

Db 153 PPCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGV 212

QY 241 EVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ 300

Db 213 EVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ 272

QY 301 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDG 360

Db 273 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDG 332

QY 361 SFFLYSKLTVDKSRWQGNVFNCSVMHEALHNNHYTQKSLSPG 404

Db 333 SFFLYSKLTVDKSRWQGNVFNCSVMHEALHNNHYTQKSLSPG 376

RESULT 7

A60764

Ig gamma-3 chain C region, form LAT - human

C:Species: Homo sapiens (man)

C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 16-Jul-1999

C:Accession: A60764

R:Huck, S.; Lefranc, G.; Lefranc, M.P.

Immunogenetics 30, 250-257, 1989

A:Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 conv

A:Reference number: A60764; MUID:90007613

A:Accession: A60764

A>Status: preliminary



Biochem. J. 117, 33-47, 1970  
A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant  
A:Reference number: A90249; MUID:70207560  
A:Accession: A90249  
A:Molecule type: protein  
A:Residues: 1-30;81-326 <PIN>  
C:Genetics:  
A:Gene: GDB:IGHG4  
A:Cross-references: GDB:119340; OMIM:147130  
A:Map position: 14q32.33-14q32.33  
A:Introns: 99/1; 111/1; 221/1  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) heavy disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger disulfide bonds. Interchain (to light chain) #status experimental  
C:Superfamily: Immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IMI>  
F:99-110/Region: hinge  
F:134-203/Domain: immunoglobulin homology <IM2>  
F:240-307/Domain: immunoglobulin homology <IM3>  
F:140/Disulfide bonds: interchain (to light chain) #status experimental  
F:27-83,141-201,247-305/Disulfide bonds: #status predicted  
F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.0%; Score 1133; DB 1; Length 327;  
Best Local Similarity 92.9%; Pred. No. 1e-63; 8; Indels 0; Gaps 0;  
Matches 208; Conservative 8; Mismatches 0;

QY 181 GTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDG 240  
Db 103 GPPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDG 162

QY 241 EVNNAKTPREEQNSRYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKG 300  
Db 163 EVNNAKTPREEQNSRYRVSVLTVLHQDWLNGKEYKCKVSNKGLPSSSEKTIISKAKG 222

QY 301 PREPQVYTLPPSDELTCKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLDSDG 360  
Db 223 PREPQVYTLPPSDELTCKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLDSDG 282

QY 361 SFELYSLKTVDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPG 404  
Db 283 SFELYSLKTVDKSRWQGNVFCSSVMHEALHNHYTQKSLSLSPG 326

RESULT 10  
G3HOW1  
Ig gamma-3 heavy chain disease proteins - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1979 #sequence\_revision 23-Oct-1981 #text\_change 16-Jul-1999  
C:Accession: A90442; A92219; A90198; A93915; A02149  
R:Frangione, B.; Rosenwasser, E.; Prell, F.; Franklin, E.C.  
Biochemistry 19, 4304-4308, 1980  
A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy  
A:Reference number: A90442; MUID:81021548  
A:Contents: heavy chain disease protein Wis  
A:Accession: A90442  
A:Molecule type: protein  
A:Residues: 1-289 <FRA>  
A:Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain  
A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co  
A:Note: the sequence of residues 42-76 was taken from the reference that follows  
R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.  
J. Biol. Chem. 252, 883-889, 1977  
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication  
A:Reference number: A92219; MUID:71118561  
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W  
A:Accession: A92219  
A:Molecule type: protein  
A:Residues: 12-97 <MIC>  
A:Note: the hinge region in gamma-3 chains is about four times as long as in other gamma  
idue segment (12-28)

A:Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form in  
R:Wolfenstein-Todel, C.; Frangione, B.; Prell, F.; Franklin, E.C.  
Biochem. Biophys. Res. Commun. 71, 907-914, 1976  
A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of t  
A:Reference number: A90198; MUID:77021516  
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residu  
A:Accession: A90198  
A:Molecule type: protein  
A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>  
A:Note: this protein lacks most of the V region, all of the CH1 region, and part of t  
R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood,  
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982  
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deleti  
A:Reference number: A93915; MUID:82247835  
A:Contents: heavy chain disease protein Omm  
A:Accession: A93915  
A:Molecule type: mRNA  
A:Residues: 12-70;72-114;116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-  
A:Note: a carboxyl-terminal Lys is removed posttranslationally  
A:Note: this sequence may represent an allelic form or another gamma chain subclass  
C:Comment: The heavy chain disease protein Wis is shown.  
C:Genetics:  
A:Gene: GDB:IGHG3  
A:Cross-references: GDB:119339; OMIM:147120  
A:Map position: 14q32.33-14q32.33  
C:Superfamily: Immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid  
F:203-270/Domain: immunoglobulin homology <IMM>  
F:1/Modified site: pyroglutamic acid (Gln) #status experimental  
F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 49.7%; Score 1125, 5; DB 1; Length 289;  
Best Local Similarity 69.6%; Pred. No. 2.7e-63; 19; Mismatches 23; Indels 53; Gaps 5;  
Matches 217; Conservative 19;

QY 97 CLKHRSCTPPPPGFGVYVAGTPERTVTCRCPCDGGFFNETSSKAP-CRKHTNCSVFGLL 152  
Db 27 CPERKSCDTPPP-----CPRCPEP-KSCDTPPPCPCPEKSCDT----- 65

QY 153 TQGNATHDNICSGNSESTOKSGGGGGGTCPCPCAPPELLGGPSVFLFPPKPKDTLMIS 212  
Db 66 -----PPCPCPCAPPELLGGPSVFLFPPKPKDTLMIS 97

QY 213 RPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQNSRYRVSVLTVLHQDWL 272  
Db 98 RPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQNSRYRVSVLTVLHQDWL 157

QY 273 NGKEYCKVSNKALPAPIEKTIISKAKGPREPQVYTLPPSDELTCKNOVSLTCLVKGFY 332  
Db 158 DGKEYCKVSNKALPAPIEKTIISKAKGPREPQVYTLPPSDELTCKNOVSLTCLVKGFY 217

QY 333 SDIAVWESNGOPENNYKTTTPVLDSDGSGFFLYSLKTVDKSRWQGNVFCSSVMHEALHN 392  
Db 218 SDIAVWESNGOPENNYKTTTPVLDSDGSGFFLYSLKTVDKSRWQGNVFCSSVMHEALHN 277

QY 393 HYTKSLSLSPG 404  
Db 278 RFTQKSLSLSPG 289

RESULT 11  
GHRB  
Ig gamma chain C region - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 24-Apr-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Jul-1999  
C:Accession: A91749; A90248; A93928; A90245; A94416; A02161  
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.  
Immunogenetics 18, 387-397, 1983  
A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I hsp  
A:Reference number: A91749; MUID:84030930  
A:Accession: A91749  
A:Molecule type: mRNA

A;Residues: 1-323 <BER>  
A;Note: This sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-Thr  
R;Pratt, D.M.; Mole, L.E.  
Biochem. J. 151, 337-349, 1975  
A;Title: Sequence studies on the constant region of the Fd sections of rabbit immunoglobulin  
A;Reference number: A90290; MUID:76135469  
A;Accession: A90290  
A;Molecule type: protein  
A;Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>  
R;Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
A;Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy chain  
A;Reference number: A93928; MUID:83299917  
A;Accession: A93928  
A;Molecule type: mRNA  
A;Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>  
A;Cross-references: GB:M16426; NID:g165111; PID:AAA31289.1; PID:g165112  
A;Note: This sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic marker  
R;Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.  
Biochem. J. 116, 249-259, 1970  
A;Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin  
A;Reference number: A90245; MUID:70110015  
A;Accession: A90245  
A;Molecule type: protein  
A;Residues: 132-143, 'E', 145-161 <FRU>  
R;Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
In Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almquist and Wiksell.  
A;Reference number: A94416  
A;Accession: A94416  
A;Molecule type: protein  
A;Residues: 129-131;155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, 'Q', 'Q'  
A;Note: This has the e15 allotypic marker, 185-Ala  
C;Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger multimers.  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin  
F:130-199/Domain: immunoglobulin homology <IM1>  
F:20-82/Domain: immunoglobulin homology <IM2>  
F:236-303/Domain: immunoglobulin homology <IM3>  
F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.2%; Score 909; DB 1; Length 323;  
Best Local Similarity 72.8%; Pred. No. 8.9e-50;  
Matches 163; Conservative 28; Mismatches 31; Indels 2; Gaps 1;  
A;Reference number: 147159; MUID:95015845

QY 183 TC-PPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGV 240  
DB 99 TCSKPTCPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYDGV 158  
QY 241 EVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAKGQ 300  
DB 159 QVTRAPRLERQGFNSTIRVSVLPITQHDWLNKGEKCKVSNKALPAPIETISKAKGQ 218  
QY 301 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDSDG 360  
DB 219 PLEPKVYTMGPPELSSRSYSLTCMINGFYPDSISVEWERNKRAEDNYKTPPAVLDSG 278  
QY 361 SFFLYSKLTVDKSRWQQGNVSCVMHEALHNYTKQSLSPG 404  
DB 279 SYFLYKLSVPTSEWQGRDVFCTCSVMHEALHNYTKQSLSPG 322

RESULT 12  
147160  
Ig gamma 2b chain constant region - pig (fragment)  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
A;Reference number: 147158; MUID:95015845

A;Accession: 147160  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-328 <KAC>  
A;Cross-references: EMBL:U03780; NID:g433125; PIDN:AAA52218.1; PID:g433126  
C;Genetics:  
A;Gene: IgG2b  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IM>

Query Match 39.9%; Score 902.5; DB 2; Length 328;  
Best Local Similarity 59.7%; Pred. No. 2.3e-49;  
Matches 172; Conservative 36; Mismatches 53; Indels 27; Gaps 4;  
A;Reference number: 147158; MUID:95015845

QY 125 PDGFFSNETSCKAPCRKHTNCSVFGLLLTQGNATHDNCISGSESTOKSGGGGGGTC 184  
DB 59 PSGLYSLSSVMYVFPASSLSKSY-----TCNVNHPATTTKDKRVRGTTK 103  
QY 185 PPCP-----APELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 238  
DB 104 PPCPICPACESP-----GPSVFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 159  
QY 239 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 298  
DB 160 GVEVHTAQTREKPEQFNSTYRVVSVLPITQHDWLNKGEKCKVSNKALPAPIETISKAK 219  
QY 299 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQ--PENNYKTPPPV 356  
DB 220 GQTPREPQVYTLPPHAELSRKSVITCLVIGYPPDIDVWQRNGQPEPEGYRTTPPQ 279  
QY 357 DSDGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNYTKQSLSPG 404  
DB 280 DVDGTIFLYKFSVDKASWQGGIFQCAVMHEALHNYTKQSLKTPG 327

RESULT 13  
147159  
Ig gamma 2a chain constant region - pig (fragment)  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 21-Jan-2000  
A;Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a  
A;Reference number: 147158; MUID:95015845

A;Accession: 147159  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-328 <KAC>  
A;Cross-references: EMBL:U03779; NID:g433123; PIDN:AAA52217.1; PID:g433124  
C;Genetics:  
A;Gene: IgG2a  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
F:133-202/Domain: immunoglobulin homology <IM>

Query Match 39.9%; Score 902.5; DB 2; Length 328;  
Best Local Similarity 59.7%; Pred. No. 2.3e-49;  
Matches 172; Conservative 36; Mismatches 53; Indels 27; Gaps 4;  
A;Reference number: 147158; MUID:95015845

QY 125 PDGFFSNETSCKAPCRKHTNCSVFGLLLTQGNATHDNCISGSESTOKSGGGGGGTC 184  
DB 59 PSGLYSLSSVMYVFPASSLSKSY-----TCNVNHPATTTKDKRVRGTTK 103  
QY 185 PPCP-----APELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 238  
DB 104 PPCPICPACESP-----GPSVFIFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD 159  
QY 239 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 298  
DB 160 GVEVHTAQTREKPEQFNSTYRVVSVLPITQHDWLNKGEKCKVSNKALPAPIETISKAK 219

A;Molecule type: protein

A;Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. Antigenic determinants of the heavy chain.  
A;Reference number: A90352; MUID:71058471



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:06 ; Search time 70.93 Seconds  
(without alignments)  
194.760 Million cell updates/sec

Title: US-09-389-782A-7

Perfect score: 2264

Sequence: 1 ETFFPKYLHYDEETSHQLC.....VMHEALHNHYTKSLSPG 404

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_36.\*

1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1467	64.8	518	15 R51003	Sequence of a reco
2	1460	64.5	485	13 R24016	Fusion protein TNF
3	1332	58.8	376	19 W60037	Antigenic peptide
4	1317	58.2	376	18 W50287	Human Fas antigen
5	1299	57.4	438	16 R81882	Plasmid pBC406/OX4
6	1299	57.4	438	19 W48976	Ox40/Fc mutein. C
7	1253	55.3	764	21 Y68949	Fusion protein of
8	1227.5	54.2	375	18 W06683	Heregulin-alpha fu
9	1226	54.2	535	20 Y17414	SVPHI-26 disintegr
10	1225	54.1	963	19 W70540	Integrin beta-1 ch
11	1223.5	54.0	664	19 W71603	Human neurturin re
12	1223.5	54.0	664	21 Y80123	Human NTNR alpha a

13	1221.5	54.0	664	19 W71604	Rat neurturin rece
14	1221.5	54.0	664	21 Y80124	Rat NTNR alpha and
15	1220.5	53.9	400	21 Y15123	Porcine CTLA-4-Ig
16	1217.5	53.8	449	20 W92678	Human noggin/immun
17	1216.5	53.7	388	19 W73513	Rabbit TGFbetaRII:
18	1216.5	53.7	388	21 Y54063	Amino acid sequenc
19	1216	53.7	784	21 Y92207	IL-13/IL-4 dual tr
20	1215.5	53.7	423	21 Y70869	Human interferon-b
21	1215.5	53.7	1158	21 Y92205	Fusion polypeptide
22	1215.5	53.7	1168	21 Y92204	Fusion polypeptide
23	1215	53.7	445	20 Y24153	Bovine LOX-1 extra
24	1214	53.6	592	20 W07097	Human Interleukin-
25	1214	53.6	592	21 Y92185	Human IL-6R-alpha-
26	1212.5	53.6	691	21 Y92202	Fusion polypeptide
27	1212.5	53.6	694	21 Y92203	Fusion polypeptide
28	1212	53.5	388	19 W73514	Human TGFbetaRII:F
29	1212	53.5	388	21 Y54064	Amino acid sequenc
30	1210.5	53.5	387	17 R90920	IL4.Y124D/IgG1 pro
31	1210.5	53.5	408	21 Y44461	Human Interleukin
32	1207.5	53.3	680	15 R48037	tICAM(453)IgG immu
33	1207.5	53.3	859	20 W70796	Human gp130-Fc-His
34	1207.5	53.3	859	21 Y92184	Human gp130-Fc-His
35	1207	53.3	595	20 W86003	Anti-574 single ch
36	1206	53.3	396	18 W18574	Aggrecanase artifi
37	1206	53.3	396	18 W18575	Aggrecanase artifi
38	1205.5	53.2	471	21 Y45030	HUMAN OCR10-Fc fus
39	1204.5	53.2	482	19 W31646	Human cytokine rec
40	1204	53.2	232	18 W26232	Human IgG1 hinge/F
41	1204	53.2	233	20 Y06617	Human Fc (IgG1).
42	1204	53.2	235	20 Y01372	Amino acid sequenc
43	1204	53.2	254	17 R89441	IgG1 hinge, CH2 an
44	1204	53.2	259	20 Y24154	Protein from pCd51
45	1204	53.2	329	17 R91806	Human immunoglobul

#### ALIGNMENTS

##### RESULT 1

R51003  
ID R51003 standard; Protein; 518 AA.

XX R51003;

XX R51003;

DT 07-OCT-1994 (first entry)

DE Sequence of a recombinant human (rhu) tumour necrosis factor receptor

DE TNFR/fc fusion protein.

XX Tumour necrosis factor receptor; chimeric antibody molecule;

KW immunoglobulin.

XX Synthetic.

XX WO9406476-A.

PN 31-MAR-1994.

XX 14-SEP-1993; 93WO-US08666.

XX 15-SEP-1992; 92US-0946236.

XX (IMMV ) IMMUNEX CORP.

XX Jacobs CA, Smith CA;

XX WPI; 1994-118172/14.

XX N-PSDB; Q45225.

PT Treating TNF mediated inflammatory diseases with TNF antagonist -  
PT esp. soluble form of TNF receptor, opt. as fusion protein with  
PT human immunoglobulin Fc region, esp. for treating arthritis





|||||  
 Db 339 tyrvsvltvlhg-wlmgkeyckvsnkalpapiextiskakgpprepvytlppsrdel 397  
 QY 317 TKNQVSLTCLVKGFPYSDIAVENSGQPNYKTTTPPVLDSDGSPFLYSLKLVVDKSRWQ 376  
 Db 398 tknqvslclvkgfypsdlavvesngqpennykttppvldsdgsfl-skltvdksrwg 456  
 QY 377 QGNVFCVSMHEALHNHYTKLSLSPG 404  
 Db 457 qgnvfscvsmhealhnhytqkslspsg 484

# RESULT 3

W60037

ID W60037 standard; Protein; 376 AA.

XX AC W60037;

XX DT 11-SEP-1998 (first entry)

XX DE Antigenic peptide hFas (nd29) containing FC region.

XX KW Fas ligand; Fas antagonist; apoptosis related disease; liver disease;  
 KW heart failure; kidney failure; graft-versus-host disease; antibody;  
 KW myocardial infarction; ischemic restenosis; endotoxemic shock.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Peptide 1..16  
 FT Protein /note= "hFas antigen signal peptide"  
 FT /note= "hFas (nd29) protein"

XX PN W09818487-A1.

XX PD 07-MAY-1998.

XX PF 31-OCT-1997; 97WO-JP03978.

XX PR 26-SEP-1997; 97JP-0262521.

XX PR 31-OCT-1996; 96JP-0290459.

XX PR 27-DEC-1996; 96JP-0351718.

XX PA (MOCH ) MOCHIDA PHARM CO LTD.

XX PA (OSAB-) OSAKA BIOSCIENCE INST.

XX PI Nagata S, Suda T, Yatomi T;

XX DR WPI; 1998-271925/24.

XX DR N-PSDB; V34430.

XX XX Use of Fas antagonist for treatment and prevention of

PT apoptosis-related diseases - such as heart or kidney failure,

PT graft-versus-host disease or liver disease

XX PS Examples; Fig 5-9; 86pp; Japanese.

XX CC This represents the antigenic peptide hFas (nd29) containing the FC  
 CC region. The invention provides the use of Fas antagonist as an agent for  
 CC the treatment and prevention of apoptosis-related diseases. The Fas  
 CC antagonist can be a partial Fas antigen peptide containing the  
 CC extracellular part of the protein, but lacking the signal sequence, an  
 CC anti-Fas antibody, or an anti-Fas ligand antibody, where the antibody is  
 CC preferably a humanised antibody. The Fas antagonist is used in the  
 CC treatment and prevention of diseases such as myocardial infarction, heart  
 CC failure, ischemic heart disease, acute kidney failure, graft-versus-host  
 CC disease, ischemic restenosis of the heart, liver or kidney, and  
 CC endotoxemic shock, and also as an organ preservative in transplantation.  
 CC The agent is of low toxicity but effectively inhibits the Fas/Fas ligand  
 CC system.

XX XX Sequence 376 AA;

Query Match 58.8%; Score 1332; DB 19; Length 376;

Best Local Similarity 66.7%; Pred. No. 4.3e-77;

Matches 270; Conservative 15; Mismatches 60; Indels 60; Gaps 11;

QY 8 LHYDEETSHQLLCKDCKPPGTYLKHQCTAKW-KTVCAPCPD-HYVTDVSWHTSDCLYCSPV 65

Db 23 lhhdgqfchk----pcppgerkardctvngdepdcvpcqegkeytdkahfsskrcrrc-l 77

QY 66 CKELQYVKOE--CNRTHNRVCECKEGRYLEIEFCLKHSRCP-----GFGVVQAGTPERNT 119

Db 78 cdegghgleveinctrtqntkcrckpnffcnstvc---ehcdpctkcehglikectltsnt 134

QY 120 VCKRCPDGFNSNETSSKAPCRKHTNCSVFGLLLLTQKNATHDNCNSENSESTQKSGGGG 179

Db 135 kcke--egrsrsnepks---cdk-----th----- 153

QY 180 GGGTCPPCAPPELLGGPSVFLFPKPKDLMISRTPEVTCVVDVSHEDPEVKFN 239

Db 154 ---tcppcapapellgpsvflfpkpkdltmistrtevtcvvvdvshedpevkfn 210

QY 240 VEVHNAKTPRERYNSTYRVVSVLVTLVHODWLNKGEYKCKVSNKALPAPIERTISKAG 299

Db 211 vevhnaaktprerqynstyrsvsvltvlhqdwlngkeyckckvsnkalpapiertiskag 270

QY 300 QPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVENSGQPNYKTTTPPVLDSD 359

Db 271 qprepqvtytlppsrdeltnqvslclvkgfypsdlavvesngqpennykttppvldsd 330

QY 360 GSPFLYSLKLTVDKSRWQOGNVFSCVSMHEALHNHYTKLSLSPG 404

Db 331 gsfllyskltvdksrwqggnvfscvsmhealhnhytqkslspsg 375

## RESULT 4

W50287

ID W50287 standard; Protein; 376 AA.

XX AC W50287;

XX DT 16-JUL-1998 (first entry)

XX DE Human Fas antigen derivative/IgG1 FC fusion.

XX KW Human; Fas antigen; derivative; apoptosis regulation; gene therapy;  
 KW treatment; diabetes; arthritis; lupus; hepatitis; influenza; HIV;  
 KW apoptosis modulation; immunoglobulin G1 FC; IgG1 FC; fusion.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..16

FT Peptide /label= sig\_peptide

FT Peptide 17..376

FT Peptide /label= mat\_peptide

XX PN W09742319-A1.

XX PD 13-NOV-1997.

XX PF 01-MAY-1997; 97WO-JP01502.

XX PR 02-MAY-1996; 96JP-0135760.

XX PA (MOCH ) MOCHIDA PHARM CO LTD.

XX PA (OSAB-) OSAKA BIOSCIENCE INST.

XX PI Nagata S, Nakamura N;

XX DR WPI; 1997-558981/51.

XX DR N-PSDB; V07004.

XX XX

PT Fas antigen derivative containing modified extracellular region -  
PT has low antigenicity, promotes apoptosis and is useful in treatment  
XX of viral and other diseases  
PS Disclosure; Fig 4; 102pp; Japanese.  
XX  
CC The present sequence is a Fas antigen derivative/IgG1 Fc  
CC fusion, which contains a Fas antigen extracellular region lacking  
CC one or more amino acid residues in the region from the  
CC amino-terminal to (but excluding) the 1st cysteine residue  
CC (preferably at least 29 residues are deleted).  
CC The derivative is an effective regulator of apoptosis and can be  
CC used (either by administration of the polypeptide, or by the use  
CC of the coding DNA in gene therapy) to treat a range of diseases,  
CC e.g. diabetes, arthritis, lupus and in particular viral diseases  
CC such as hepatitis, influenza and HIV, by modulating apoptosis of  
CC virus-infected cells.  
XX  
SQ Sequence 376 AA;  
  
Query Match 58.28; Score 1317; DB 18; Length 376;  
Best Local Similarity 65.94; Pred. No. 3.9e-76;  
Matches 267; Conservative 16; Mismatches 62; Indels 60; Gaps 11;  
  
QY 8 LHYDEETSHQLLCKPPTGTLKQHTAKW-KTVCAPCPD-HYITDSWHTSDECLYCSVP 65  
DB 23 lhhdqfchk----pcpggarkardctvngdepdcvpcqeqkeytdkahfskrrcr-l 77  
  
QY 66 CKELQYVQOE--CNRTHNRVCEKRGYLEIEFCLKHSRCPD---GFGVVOAGTPERNT 119  
DB 78 cdeghgleveinctrtgtktcrckpnffcnstvc---ehcdpctkcehglikectltst 134  
  
QY 120 VKCRCPDGFNSETSKAPCRKHTNCNVFGLLLTQKGNATHDNCGNSSTQKSGGGG 179  
DB 135 kcke--egsrnepks---cdk-----th----- 153  
  
QY 180 GGGTCCPAPPELLGSPVFLFPKPKDTLMISRTPEVTVVVDVSHEDPEVKFNWYVDG 239  
DB 154 ---tcpcpapellggpsvflfpkpkdtlmisrptevtcvvdvshedpevkfnwyvdg 210  
  
QY 240 VEVHNAKTPREEQYNSTYRVSVTLVHQDLNGKEYCKVSNKALPAPIEKTISKAG 299  
DB 211 vevhnaaktprceegynstyrsvsvtlvldhngkeyckvsnkalpapiektiskag 270  
  
QY 300 QPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSD 359  
DB 271 qpreqvylppsrdeltknqvsltcclvkgyfypsdiavewesngqpennkttppvldsd 330  
  
QY 360 GSFFLYSLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSPG 404  
DB 331 gsfflysltvdksrwqgnvfscsvmhlealhnhytqkslspsg 375  
  
RESULT 5  
ID R81882  
XX R81882 standard; Protein; 438 AA.  
AC R81882;  
XX  
XX 30-MAR-1996 (first entry)  
XX  
XX Plasmid pBC406/OX40/Fc\* encoding an OX40/Fc mutein protein.  
XX  
XX OX40; OX40-L; cytokine; cell surface molecule; plasmid;  
KW pDC406/OX40/Fc\*; membrane glycoprotein.  
XX  
XX Synthetic.  
XX  
XX US5457035-A.  
XX  
XX 10-OCT-1995.  
XX

PF 23-JUL-1993; 93US-0097827.  
XX  
XX 23-JUL-1993; 93US-0097827.  
XX  
XX (IMMV ) IMMUNEX CORP.  
XX  
XX Baum PR, Fanslow WC, Gayle RB, Goodwin RG;  
XX PI  
XX WPI: 1995-357992/46.  
DR N-PSDB; T00829.  
XX  
XX New isolated DNA encoding the OX40 ligand polypeptide - also vectors  
PT and host cells, used to produce recombinant ligand used in e.g.  
PT prim. T cell culture, to modulate immune response etc.  
XX  
XX Example 2; Column 35-38; 26pp; English.  
XX  
XX This plasmid encodes an OX40/Fc antibody fragment mutein protein,  
CC and is used to express a soluble OX40/Fc mutein fusion protein for  
CC use in detecting cDNA clones encoding a OX40 ligand. The Fc  
CC fragment may be derived from human IgG1, and the plasmid may be  
CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell  
CC line. Culture supernatant was purified by affinity chromatography  
CC and this was used, together with labeled goat anti-human IgG to  
CC screen various cell lines.  
XX  
XX Sequence 438 AA;  
  
Query Match 57.48; Score 1299; DB 16; Length 438;  
Best Local Similarity 60.44; Pred. No. 6.3e-75;  
Matches 258; Conservative 22; Mismatches 77; Indels 70; Gaps 6;  
  
QY 20 CDKCPPTGTLKQHTAKWKTVCAPCPDHYITD--SWHTSDECLYCSVPCKELQYVQECN 77  
DB 39 ccreqpgghgmvrncdtrdtclhpcetgyneavnydtckqctqcnh--rsgselkqnc 96  
  
QY 78 RTHNRVCEKRGYLEIEFCLKHSRCPDGFVQVAGTPERNTVCKRCRCPDGFSPNETSKA 137  
DB 97 ptqdtvcr-----pgtprqdsqykgldvcpvcpghfs--pgnq 137  
  
QY 138 PCRHTNCVFGLLLTQKGNATHDNCGNS----- 168  
DB 138 ackpwtntctisgktrhpasdsldavcedrsllatllwetqrptvtgqsttvprrts 197  
  
QY 169 -----ESTQKSGGGGGGTCTPCPAPPELLGGPSVFLFPKPKDTLMISRTPEV 217  
DB 198 elpatptliveprscdt-----htcppeapeaeapsvflfpkpkdtlmisrtpev 250  
  
QY 218 TCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVSVTLVHQDLNGKEY 277  
DB 251 tcvvvdvshedpevkfnwyvdgvevhnaktprceegynstyrsvsvtlvldhngkey 310  
  
QY 278 KCKVSNKALPAPIEKTISKAGQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAV 337  
DB 311 kckvsnkalpapiektiskagqpreqvylppsrdeltknqvsltcclvkgyfypsdiav 370  
  
QY 338 EWESNGOPENNYKTTTPVLDSDGSGFFLYSLTVDKSRWQGNVFCSCVMHEALHNHYTK 397  
DB 371 ewesngpennkttppvldsdgsfflysltvdksrwqgnvfscsvmhlealhnhytqk 430  
  
QY 398 SLSLSPG 404  
DB 431 slslspg 437  
  
RESULT 6  
W48976  
ID W48976 standard; Protein; 438 AA.  
XX  
XX W48976;  
AC  
XX  
XX 25-SEP-1998 (first entry)  
XX



CC within vascular endothelium indicates a role for the polypeptides  
CC in the control of endothelial cell biology. The murine polynucleotide  
CC was identified from a white adipose tissue cDNA library. The polypeptide  
CC is useful for identifying receptors, which bind to and/or are activated  
CC by the polypeptide. The polynucleotide is useful in gene therapy of  
CC cerebral autosomal dominant arteriopathy with subcortical infarcts and  
CC leucoencephalopathy, an autosomal dominant disorder causing ischemic  
CC strokes.

XX Sequence 764 AA;

Query Match 55.3%; Score 1253; DB 21; Length 764;  
Best Local Similarity 58.0%; Pred. No. 9.6e-72;  
Matches 267; Conservative 22; Mismatches 67; Indels 104; Gaps 15;

Qy 11 DEETSHQLLCKPCPGYLYLKHCHTAKWTVCA-----PCPDHY 49  
Dy 342 dqensyhcl---cpbg-yvgdqchshlt-cadspcfnngscrerngssyacecpnft 396  
Qy 50 TDSWH-----TSDECLYCSVPCKELQVYKQECNRTNHRVCECKEG-----RYLEIEFCL 98  
Dy 397 gsneckkvdrctspcang-----qclnrgpsrtcrpfgtghcelhisdca 446  
Qy 99 KHRSCPPGFGVVOAGT---PERNTVKRCPPDGF-----FSNETSKAPCRKHTNGSV 147  
Dy 447 r-spcahg-----gtchdlengpvc-tcpagfsgrrcevrithdacasgpcfnagtc-- 496  
Qy 148 FGLLLTQGNATHDNIC-----SGNSESTOKSGGGGGGTC 184  
Dy 497 -----ytlgnlfnvcncpvgfvgscfepvgippsfpaapekscdkt-----htc 543  
Qy 185 PPCPAPELLGGPSVFLRPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 244  
Dy 544 ppcpapellgpgsvflppkpkdtlmisrtpevtcvvvdshknpvfnwvydgvvevhn 603  
Qy 245 AKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREP 304  
Dy 604 aktpreeqynstyrvvsvltvlhqdwingkeyckvsnkalpapiektiskakgqprep 663  
Qy 305 QVYVLPSPRDELTKNQVSLTCLGVFPSPDIATVWESNGQPNYKTPPPVLDSDGSFFL 364  
Dy 664 qvytlpsrdeltknqvsitclvkgfypsdiavewesngqpennyyktpvldsdgsffl 723  
Qy 365 YSKITVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 404  
Dy 724 yskitvdksrwqgnvfscvmhealhnhytcqkslsisp 763

## RESULT 8

W06683  
ID W06683 standard; Protein; 375 AA.

XX W06683;

XX 24-FEB-1997 (first entry)

DE Heregulin-alpha fusion protein HRG-alpha-T-Fc.

XX Heregulin; epidermal growth factor receptor; EGF; HER4 receptor;  
KW fusion protein; protein phosphorylation; breast cancer;  
KW rHRG-alpha-T-Fc.

OS Chimeric Homo sapiens;

OS Chimeric synthetic.

XX Key Location/Qualifiers

FT Peptide 1..59

FT /label= sig.peptide  
FT /note= "CD5 signal sequence"

FT Misc-difference 26

FT /note= "x at position 26 corresponds to an  
FT in-frame stop codon in the coding sequence"

FT Protein 60..375  
FT /label= HRG-alpha-T-Fc  
FT /note= "mature fusion protein produced following  
FT signal peptide cleavage"  
FT Region 60..124  
FT /label= HRG-alpha  
FT /note= "heregulin-alpha EGF-like binding domain"  
FT Cleavage-site 125..141  
FT /label= Thrombin\_cleavage\_site  
FT Region 142..375  
FT /label= Human Ig-constant region  
XX W09636720-A1.  
XX PN 21-NOV-1996.  
XX PD 14-MAY-1996; 96WO-US06861.  
XX PF 16-MAY-1995; 95US-0441863.  
XX PR (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX PA Aruffo AA, Carlton GW, Culouscou J;  
XX PI WPI; 1997-012095/01.  
XX DR N-PSDB; T45262.  
XX Eukaryotic vector encoding growth factor, thrombin cleavage site and  
PT human IgG1 Fc region - esp. encoding EGF-like domain of  
PT heregulin-alpha, -beta3, used for prodn. of recombinant  
PT fusion proteins capable of HER4 receptor activation  
XX Example 1; Fig 7A-B; 188pp; English.  
XX rHRG-alpha-T-Fc (W06683) is a fusion protein that incorporates the  
CC EGF-like binding domain of human heregulin (HRG)-alpha and the Fc  
CC portion of human IgG1. It is the product of a cDNA construct  
CC (T45262) in vector CDW7 and can be produced in eukaryotic (esp.  
CC COS) host cells. The fusion protein, or the HRG-alpha EGF binding  
CC domain produced from it by thrombin cleavage, are both capable of  
CC activating human epidermal growth factor receptor 4 (HER4) and can  
CC be used to study receptor functioning, e.g. protein phosphorylation.  
CC They are also useful in screening assays for (ant)agonists of HER3  
CC and HER4, and bind to cells that express HER4 e.g. MDA-MB-453 human  
CC breast cancer cells, inducing expression of intercellular adhesion  
CC molecule-1.  
XX Sequence 375 AA;

Query Match 54.2%; Score 1227.5; DB 18; Length 375;  
Best Local Similarity 69.9%; Pred. No. 1.8e-70;  
Matches 246; Conservative 7; Mismatches 44; Indels 55; Gaps 7;

Qy 76 CNRTHRVCECKEGRYLEIEFCLKHSRCPGFGVVGQAGTPERTVCKRCPDGFSSNETSS 135  
Dy 55 clgtsshlvkae---kectfcvnggec---fmvkdlnpsr-ylck-cqpgfigarcte 106  
Qy 136 KAPCRKHTNCSVFGLLLTQGNATHDNICSNSEST-----OKSGGGGGG----- 180  
Dy 107 nvpmkvq-----nqekaeglyqkdp9ggggrivprfgsgd 142  
Qy 181 -----GGTCPPCPAPELLGGPSVFLRPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 232  
Dy 143 pepkscdkthccppcpapefegapsvflppkpkdtlmisrtpevtcvvvdshedpevk 202  
Qy 233 FNWYVDGVEVHNKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 292  
Dy 203 fnwydgvvevhnaktpreeqynstyrvvsvltvlhqdwingkeyckvsnkalpapi 262  
Qy 293 TISKAKGQPREPQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 352  
Dy 263 tiskakgqprepyvtlppsrdeitknqvsitclvkgfypsdiavewesngqpennyykt 322

QY 353 PPVLDSGSGFFLYSKLTVKSRWQGNVFCSCVMHEALHNHYTKQKSLSPG 404  
Db 323 ppvltdsgsflyskltvdkrsrwwgnvfcscvmhealhnhytqkslsisp 374

RESULT 9  
Y17414  
ID Y17414 standard; Protein; 535 AA.  
AC Y17414;  
XX  
XX 26-JUL-1999 (first entry)  
XX  
XX SVPH1-26 disintegrin domain, Ig kappa signal, Fc IgG1 fusion protein.  
DE Human; SVPH1-26; proteinase; testis; fertilisation; spermatogenesis;  
KW birth control; detergent additive; diagnosis; testicular cancer.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX Location/Qualifiers  
FH 1..20  
FT /note= "Ig kappa signal sequence"  
FT 21..22  
FT /label= spacer  
FT 23..305  
FT /note= "SVPH1-26 disintegrin domain"  
FT 306..307  
FT /label= spacer  
FT 308..535  
FT /note= "IgG1 Fc domain"  
XX  
XX WO9223228-A1.  
XX  
XX 14-MAY-1999.  
XX  
XX 30-OCT-1998; 98WO-US22965.  
XX  
XX 30-OCT-1997; 97US-0063571.  
XX  
XX (IMMV ) IMMUNEX CORP.  
XX  
XX Cerretti DP;  
XX  
XX WPI; 1999-337487/28.  
XX  
XX New human SVPH1-26 DNA useful for the diagnosis and prognosis of  
XX testicular cancers  
XX  
XX Example 2; Page 59; 96pp; English.  
XX  
XX The present invention describes human SVPH1-26. Human SVPH1-26  
XX polypeptides are proteinases implicated in fertilisation and  
XX spermatogenesis. They can be used as therapeutic agents. A proteinase  
XX inhibitor of the catalytic domain would inhibit SVPH1-26 activity and  
XX would be useful as a method for birth control. Also, an inhibitor of the  
XX disintegrin domain of SVPH1-26 may affect fertilisation. The proteinase  
XX activity of SVPH1-26 can also be used as a detergent additive for the  
XX removal of stains having a protein component. The SVPH1-26 polypeptides  
XX and fragments can also be used as molecular weight markers, as markers  
XX for determination of isoelectric points of sample proteins, and as  
XX controls for establishing the extent of fragmentation of a protein  
XX sample. The products can also be used for identifying, separating or  
XX purifying cells that express SVPH1-26 polypeptides such as testis cells.  
XX They can be used for the diagnosis and prognosis of testicular cancers.  
XX The present sequence represents a fusion protein from an example of the  
XX present invention comprising: the Ig kappa signal sequence at positions  
XX 1 to 20; a spacer at positions 21 to 22; the SVPH1-26 disintegrin domain  
XX at positions 23 to 305; a spacer at positions 306 to 307; and the  
XX IgG1 Fc domain at positions 308 to 535.

SQ Sequence 535 AA;

Query Match 54.2%; Score 1226; DB 20; Length 535;  
Best Local Similarity 54.3%; Pred. No. 3.3e-70;  
Matches 268; Conservative 18; Mismatches 98; Indels 110; Gaps 14;

QY 8 LHYDEETSHOLLCDKC---PPGTYLKQ---HC-TAKW-KTVCAPCPDPHYTDSWHTSDEC 59  
Db 54 lhpgaacafgicckdckflpsgtlcrqgvedclpewcngtshqcpddvvyqdgiscvn 113  
QY 60 LYC-----SPVKCEL-----QYVKQECNTHNRV---CECKEGRYLEI----- 94  
Db 114 afcyekctnhdhlgckefgqdsarsasqcyqelntqgnrfghcgivgttyvkcwtpdim 173  
QY 95 -----EFCLKHRSC-----PPGFGVVOAGT---PERNT 119  
Db 174 cgrvqcenvgvipnllehtstvgqfhdtttcwgtgdyhlgmaipdigevkdgvtcgpellic 233  
QY 120 VCKRCPDGFNETSSKAPCRKHTNCVSFGLLLLTQGNATHDNICSGNSE-----STQKS 174  
Db 234 irkkcasmvhlsgacqpkctcnmrgicn-----nkqh---chcnhewappycdkg 280  
QY 175 GGGGGGGG-----TCPPCPAPELLGGPSVFLFPKPKDPTLM 210  
Db 281 ygsadsqppknmeglnvmgklrgscdkthtccpcpapeaegapsvflfpkpkdtlm 340  
QY 211 ISRTPEVTCVVVDYSHEDPEVKFNWYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHQD 270  
Db 341 isrtpevtcvvdshedpevkfnwyvdgvevhnaktkpreeqynstyrsvsvltvlhqd 400  
QY 271 WLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGF 330  
Db 401 wlngkeyckvsnkalpapiektiskakgqprepyvltlppsrdeltknqvslclvkgf 460  
QY 331 YPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSLKLTVDKSRWQGNVFCSCVMHEAL 390  
Db 461 ypsdiavewesngqpennnykttppvldsdgsfflyskltvdkrsrwwqgnvfscsvmheal 520  
QY 391 HNHYTKSLSPG 404  
Db 521 hnhytqkslsisp 534

RESULT 10

W70540  
ID W70540 standard; Protein; 963 AA.

XX AC W70540;

XX DT 26-JAN-1999 (first entry)

XX DE Integrin beta-1 chain.

XX KW Integrin; beta-1 chain; immunoglobulin; chimeric; heterodimer complex;  
XX KW inhibitor; binding; ligand; blood platelet; hemostatic; diagnostic agent;  
XX KW human.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Peptide 1..20 /note= "signal peptide"

XX FT Protein 21..963

XX FT /note= "mature protein"

XX PN WO9832771-A1.

XX PD 30-JUL-1998.

XX PF 29-JAN-1998; 98WO-JP00370.

XX PR 29-AUG-1997; 97JP-0234544.



Db 274 ncrasyqvtls-cpadnyqacclgsyagimigdmtpnyvdsptg---ivvspwc----- 323  
 QY 73 QKCNRTNRHVCECKEGRYLEIEF---CLKH-----RSCP--PGFGVVOAGTP 115  
 Db 324 --scrgsgmneeece--kfllr-dftenpcrlnaiafngtdvnsvpkpsfqtatqprv 378  
 QY 116 ERNTVCKRCPDGPFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHD-----NICS 165  
 Db 379 ekt---pslpddlsdstslgtstvtctsvqeggl---kannskelsmofteitnliip 431  
 QY 166 GNSESTQKSGGGGGGTCTPCPAPPELLGSPSVFLFPKPKDTLMISRTPEVTCVVVDVS 225  
 Db 432 gprdpvdk-----htcpcpapellggsvflfppkpkdtlmisrtpevtcvvvdvs 484  
 QY 226 HEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVVSVLTVLHODWLNGLNGKEYCKVSNKA 285  
 Db 485 hedpevkfnwyvvgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeyckvsnka 544  
 QY 286 LPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVRGFYPSDIAVEWESNGQP 345  
 Db 545 lpapiektiskakgqprepqvyltlppsreemtknqsltlclvkgyfypsdiavewesngqp 604  
 QY 346 ENNYKTTTPPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKSLSPG 404  
 Db 605 ennykttppvldsdgsfflyskltvdksrwqgnvfscsvmhealhnyhtqkslisp 663

RESULT 12  
 Y80123  
 ID Y80123 standard; Protein; 664 AA.  
 XX  
 AC Y80123;  
 XX  
 DT 19-MAY-2000 (first entry)  
 XX  
 DE Human NTNR alpha and IgG fusion protein SEQ ID NO:16.  
 XX  
 KW Human; neuturin receptor alpha; NTNR alpha; splenic haematopoiesis;  
 KW anaemia; thrombocytopaenia; hypoplasia; haemorrhage.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN US6025157-A.  
 XX  
 PD 15-FEB-2000.  
 XX  
 PF 24-OCT-1997; 97US-0957063.  
 XX  
 PR 18-FEB-1997; 97US-0038839.  
 PR 09-JUN-1997; 97US-0049818.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Hynes MA, Rosenthal A, Klein RD;  
 XX  
 DR WPI: 2000-181808/16.  
 DR N-PSDB; 291459.  
 XX  
 PT Isolated nucleic acid molecule encodes a neuturin receptor-alpha amino  
 PT acid sequence excluding the N-terminal signal peptide -  
 XX  
 PS Claim 3; Column 91-96; 78pp; English.  
 XX  
 CC The present invention describes a neuturin receptor alpha (NTNR alpha).  
 CC NTNR alpha binds neuturin. The NTNR alpha nucleic acid molecule is  
 CC useful for the expression of NTNR alpha, which may be used to identify  
 CC agonist and antagonist compounds having therapeutic applications, such  
 CC as enhancing splenic haematopoiesis, treating anaemia,  
 CC thrombocytopaenia, hypoplasia, or haemorrhage. The present sequence  
 CC is a NTNR alpha and IgG (immunoglobulin G) fusion protein from the  
 CC present invention.  
 XX

SQ Sequence 664 AA;  
 Query Match 54.0%; Score 1223.5; DB 21; Length 664  
 Best Local Similarity 62.1%; Pred. No. 6e-70;  
 Matches 260; Conservative 24; Mismatches 60; Indels 75; Gaps 14;  
 QY 32 HCTAKWKTVCAPCP-DHY-----YTDSWHTSDEGLYCSPVCKELQYV 72  
 Db 274 ncrasyqvtls-cpadnyqacclgsyagimigdmtpnyvdsptg---ivvspwc----- 323  
 QY 73 QKCNRTNRHVCECKEGRYLEIEF---CLKH-----RSCP--PGFGVVOAGTP 115  
 Db 324 --scrgsgmneeece--kfllr-dftenpcrlnaiafngtdvnsvpkpsfqtatqprv 378  
 QY 116 ERNTVCKRCPDGPFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHD-----NICS 165  
 Db 379 ekt---pslpddlsdstslgtstvtctsvqeggl---kannskelsmofteitnliip 431  
 QY 166 GNSESTQKSGGGGGGTCTPCPAPPELLGSPSVFLFPKPKDTLMISRTPEVTCVVVDVS 225  
 Db 432 gprdpvdk-----htcpcpapellggsvflfppkpkdtlmisrtpevtcvvvdvs 484  
 QY 226 HEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVVSVLTVLHODWLNGLNGKEYCKVSNKA 285  
 Db 485 hedpevkfnwyvvgvevhnaktprreeqynstyrvvsvltvlhqdwlngkeyckvsnka 544  
 QY 286 LPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVRGFYPSDIAVEWESNGQP 345  
 Db 545 lpapiektiskakgqprepqvyltlppsreemtknqsltlclvkgyfypsdiavewesngqp 604  
 QY 346 ENNYKTTTPPVLDSDGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKSLSPG 404  
 Db 605 ennykttppvldsdgsfflyskltvdksrwqgnvfscsvmhealhnyhtqkslisp 663

RESULT 13  
 W71604  
 ID W71604 standard; Protein; 664 AA.  
 XX  
 AC W71604;  
 XX  
 DT 19-NOV-1998 (first entry)  
 XX  
 DE Rat neuturin receptor alpha/Fc sequence (If2a) fusion protein.  
 XX  
 KW Human; neuturin receptor alpha; NTNR-alpha; variant; chimeric;  
 KW fusion protein; immunoadhesion; ret-expressing cell; neurological;  
 KW renal; haematological disease.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 OS Rattus sp.  
 XX  
 PN WO9836072-A1.  
 XX  
 PD 20-AUG-1998.  
 XX  
 PF 17-FEB-1998; 98WO-US03179.  
 XX  
 PR 24-OCT-1997; 97US-0957063.  
 PR 18-FEB-1997; 97US-0802805.  
 PR 09-JUN-1997; 97US-0871913.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Hynes MA, Klein RD, Rosenthal A;  
 XX  
 XX WPI: 1998-467175/40.  
 DR N-PSDB; V58007.  
 XX  
 PT New polypeptide(s) based on human neuturin receptor alpha and  
 PT related nucleic acid - useful for increasing survival of





```
AC Y15123;
XX
XX 07-FEB-2000 (first entry)
XX
XX Porcine CTLA-4-Ig construct.
XX
XX Porcine CTLA-4; soluble protein; xenograft; organ transplant; B7; CD28;
KW xenograft-specific immunosuppression; recipient T-cell; anergy;
KW co-stimulatory signal 2; homology; human CTLA-4; bovine CTLA-4.
XX
XX Sus scrofa.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX 162..168
FT Region /label="Flexible_linker
FT /note="Denotes the junction between pCTLA-4"
FT Domain 169..362
FT /label="IgG1_domain
FT Domain
XX
XX WO957266-A2.
XX
XX 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-GB01350.
XX
XX 30-APR-1998; 98GB-0009280.
XX
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
XX Lechler IR, Dorling A;
XX
XX WPI; 2000-038815/03.
XX
XX Inhibiting T-cell mediated rejection of xenotransplanted organs
XX
XX Claim 1; Fig 4; 43pp; English.
XX
XX The present sequence is porcine CTLA-4-Ig construct for xenograft
XX -specific immunosuppression. In a pig-to-human transplantation, the
XX soluble protein could comprise the extracellular domain of porcine CTLA-4
XX fused to a human C gamma 1 chain of IgG1. This construct was subcloned
XX into the expression vector pHOOK-3TM and used to transfect DAP.3 or
XX CHO-K1 cells. pCTLA-4-Ig preferentially binds to porcine B7 and blocks
XX its interaction with CD28 on recipient T-cells. This is useful as a
XX species-specific reagent to inhibit human T-cell proliferative responses
XX to a variety of stimulators.
XX
XX Sequence 400 AA;

Query Match 53.9%; Score 1220.5; DB 21; Length 400;
Best Local Similarity 89.5%; Pred. No. 5.2e-70;
Matches 231; Conservative 2; Mismatches 14; Indels 11; Gaps 2;

QY 156 GNATHDNICSGNSTOKSGGGGGG-----TCPCPAPPELLGGPSVFLFPKPK 206
Db ||| : | | | | |
144 gngtqiyyi--dpepcpdsdggsgaaepkscdkthtcpcpapelligpsvflfpkpk 201
QY ||| : | | | | |
207 DTLMSIRTPETCVVVDVSHEDPEVKFNWYDGVVEHNAKTKPREOYNSTYRVVSVLT 266
Db ||| : | | | | |
202 dtlmisrtpevtcvvvdshedpevkfnwydgvvehnmaktkpreedynstyrvsvltv 261
QY ||| : | | | | |
267 LHQDLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCL 326
Db ||| : | | | | |
262 lhqdlngkeykckvsnkalpapiektiskakgprepqvylppsrdeltnqvslctl 321
QY ||| : | | | | |
327 VKGYFPDSIAVWESNGQPNNTTPPVLDSGDFLYSKLTVDKSRWQGNVFSQSV 386
Db ||| : | | | | |
322 vkgfypsdiavewesngqpennytppvltdsgdflyskltvdksrwqgnvfscsvm 381
QY ||| : | | | | |
387 HEALTHYTKSLSPG 404
```

Db 382 healthytkslspg 399

Search completed: March 1, 2001, 09:17:08  
Job time: 226 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:20:11 ; Search time 135.68 seconds  
(without alignments)  
346.406 Million cell updates/sec

Title: US-09-389-782a-8  
Perfect score: 2246  
Sequence: 1 MDKTHCTPCPAPPELLGSPS.....OKGNATHDNCISGNSESTQK 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_15:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_prodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1009	44.9	401	4 Q00300	O00300 homo sapien
2	1008	44.9	372	4 Q9UHP4	Q9uhp4 homo sapien
3	888	39.5	401	11 O08712	O08712 mus musculus
4	875	39.0	401	11 O08727	O08727 rattus norv
5	824.5	36.7	437	11 Q9RIA4	Q9ria4 mus musculus
6	433	19.3	300	4 Q95407	Q95407 homo sapien
7	405.5	18.1	302	13 Q9PUS0	Q9pus0 salvelinus
8	332.5	14.8	459	11 Q62327	Q62327 mus musculus
9	331	14.7	439	4 Q16042	Q16042 homo sapien
10	327	14.6	482	11 O08734	O08734 mus musculus
11	287	12.8	655	4 Q75509	Q75509 homo sapien
12	276	12.3	384	4 Q9UP60	Q9up60 homo sapien
13	272.5	12.1	684	13 Q90544	Q90544 ginglymosto
14	271	12.1	416	4 Q9NPP6	Q9npp6 homo sapien
15	252.5	11.2	616	4 Q9Y6Q6	Q9y6q6 homo sapien
16	252.5	11.2	625	11 O35305	O35305 mus musculus
17	234.5	10.4	348	12 O57103	O57103 monkeypox v
18	233.5	10.4	349	12 O57099	O57099 monkeypox v
19	231.5	10.3	349	12 O57291	O57291 monkeypox v

20	231.5	10.3	349	12	O57100	monkeypox v
21	231.5	10.3	349	12	O57101	monkeypox v
22	231.5	10.3	349	12	O57102	monkeypox v
23	230.5	10.3	348	12	O57277	monkeypox v
24	230.5	10.3	348	12	O57108	monkeypox v
25	229	10.2	348	12	O57112	variola vir
26	229	10.2	348	12	O85407	Q85407 variola vir
27	226.5	10.1	349	12	O57284	camelpox vi
28	226.5	10.1	349	12	O57098	camelpox vi
29	226	10.1	349	12	O57110	variola vir
30	226	10.1	349	12	O57111	variola vir
31	226	10.1	349	12	O89118	Q89118 variola vir
32	226	10.1	349	12	O89098	Q89098 variola vir
33	223	9.9	350	12	O57116	cowpox viru
34	222.5	9.9	349	12	O57097	camelpox vi
35	222.5	9.9	355	12	O85308	cowpox viru
36	221	9.8	349	12	O57109	variola vir
37	220.5	9.8	349	12	O57305	cowpox viru
38	217.5	9.7	326	12	O57120	cowpox viru
39	217.5	9.7	326	12	O57122	cowpox viru
40	217.5	9.7	351	12	O73559	cowpox viru
41	217.5	9.7	360	12	O57118	cowpox viru
42	216.5	9.6	351	12	O57117	cowpox viru
43	214.5	9.6	350	12	O57123	cowpox viru
44	212.5	9.5	347	12	O57115	cowpox viru
45	211.5	9.4	351	12	O57121	cowpox viru

## ALIGNMENTS

```

RESULT 1
O00300
ID O00300 PRELIMINARY; PRT; 401 AA.
AC O00300; O60236;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE OSTEOPTERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER IIB).
GN TNFRSF11B OR OPG OR OCIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trall G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Suggs S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
RT of bone density.".;
RL Cell 89:309-319(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG FIBROBLAST;
RX MEDLINE=98151033; PubMed=9492069;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
RA Tsuda E., Morinaga T., Higashio K.;
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
RL osteoclastogenesis in vitro.";
RL Endocrinology 139:1329-1337(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=98351569; PubMed=9688283;

```

RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;  
 RT "Cloning and characterization of the gene encoding human  
 RL Eur. J. Biochem. 254:685-691(1998).  
 CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
 CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY  
 CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
 CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
 CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
 CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,  
 CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN  
 CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN  
 CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.  
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; AB002146; BAA25910.1; -.  
 DR EMBL; AB008822; BAA32076.1; -.  
 DR EMBL; AB008821; BAA32076.1; JOINED.  
 DR HSP; U94332; AAB53709.1; -.  
 DR HSP; P25942; ICDF.  
 DR MIM; 602643; -.  
 DR INTERPRO; IPR001368; -.  
 DR PFAM; PF00020; TNFR\_C6; 3.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; PS00500; TNFR\_NGFR\_2; 2.  
 DR PRODOM; PD000771; -; 1.  
 KW Glycoprotein; Repeat; Cytokine; Signal.  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT CHAIN 22 401 OSTROPROTEGERIN.  
 FT DOMAIN 23 183 4 X TNFR-CYS.  
 FT REPEAT 23 63 TNFR-CYS 1.  
 FT REPEAT 64 106 TNFR-CYS 2.  
 FT REPEAT 107 143 TNFR-CYS 3.  
 FT REPEAT 144 201 TNFR-CYS 4.  
 FT DOMAIN 306 365 DEATH DOMAIN.  
 FT DISULFID 41 54 BY SIMILARITY.  
 FT DISULFID 44 62 BY SIMILARITY.  
 FT DISULFID 65 80 BY SIMILARITY.  
 FT DISULFID 83 97 BY SIMILARITY.  
 FT DISULFID 87 105 BY SIMILARITY.  
 FT DISULFID 118 142 BY SIMILARITY.  
 FT DISULFID 145 160 BY SIMILARITY.  
 FT CARBOHYD 98 98 POTENTIAL.  
 FT CARBOHYD 152 152 POTENTIAL.  
 FT CARBOHYD 165 165 POTENTIAL.  
 FT CARBOHYD 178 178 POTENTIAL.  
 FT CARBOHYD 289 289 POTENTIAL.  
 FT CONFLICT 263 263 A -> D (IN REF. 2 AND 3).  
 SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 44.9%; Score 1009; DB 4; Length 401;  
 Best Local Similarity 99.4%; Pred. No. 3.7e-80;  
 Matches 173; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 228 KETFPKYLHYDEETSHQLLCKDKPPGYLKQHCCTAKWTKVACPCPDHYTDSWHTSDEC 287  
 Db :|||||  
 QY 288 LYCSVPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHSRCPGPGVVOAGTTPERT 347  
 Db :|||||  
 QY 81 LYCSVPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHSRCPGPGVVOAGTTPERT 140  
 QY 348 VKCRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQK 401  
 Db :|||||  
 QY 141 VKCRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQK 194  
 RESULT 2  
 Q9UHP4 PRELIMINARY; PRT; 372 AA.

AC Q9UHP4;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE OSTROPROTEGERIN (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;  
 RT "Cloning and Expression of Osteoprotegerin from Homo sapiens";  
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).  
 DR EMBL; AF134187; AAF20168.1; -.  
 DR HSP; P25942; ICDF.  
 DR INTERPRO; IPR001368; -.  
 DR PFAM; PF00020; TNFR\_C6; 3.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; PS00500; TNFR\_NGFR\_2; 2.  
 FT NON\_TER 1  
 SQ SEQUENCE 372 AA; 42758 MW; F02527ASCD01CCD3 CRC64;  
 Query Match 44.9%; Score 1008; DB 4; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-80;  
 Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 229 ETFPKYLHYDEETSHQLLCKDKPPGYLKQHCCTAKWTKVACPCPDHYTDSWHTSDEC 288  
 Db :|||||  
 QY 289 YCSVPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHSRCPGPGVVOAGTTPERT 348  
 Db :|||||  
 QY 61 YCSVPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHSRCPGPGVVOAGTTPERT 120  
 QY 349 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQK 401  
 Db :|||||  
 QY 121 CKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQK 173  
 RESULT 3  
 O08712 PRELIMINARY; PRT; 401 AA.  
 AC O08712; 070202;  
 DT 01-JUL-1997 (Tremblrel. 04, Created)  
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE OSTROPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)  
 DE (OCIF).  
 GN TNFRSF11B OR OPG.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C; TISSUE=KIDNEY;  
 RX MEDLINE=97262071; PubMed=9108485;  
 RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,  
 Luethy R., Nguyen H.O., Woodson A., Bennett L., Boone T., Shimamoto G.,  
 Daro E., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,  
 Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,  
 Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,  
 Suggs S., Boyle W.J.;  
 RT "Osteoprotegerin: a novel secreted protein involved in the regulation  
 of bone density";  
 RL Cell 89:309-319(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/OLA, AND NIH SWISS;  
 RX MEDLINE=98382527; PubMed=9714833;

RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,  
 RA Higashio K.;  
 RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)  
 RL gene and its expression in embryogenesis.";  
 RL Gene 215:339-343(1998).  
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
 CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY  
 CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
 CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
 CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
 CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).  
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,  
 CC BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND  
 CC PLACENTA. NOT DETECTED IN SPLEEN.  
 CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT  
 CC DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY  
 CC 15 TO DAY 17.  
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; U94331; AAB53708.1; -.  
 DR EMBL; A9013898; BAA28269.1; -.  
 DR EMBL; A9013903; BAA33388.1; -.  
 DR EMBL; A9013899; BAA33388.1; JOINED.  
 DR EMBL; A9013900; BAA33388.1; JOINED.  
 DR EMBL; A9013901; BAA33388.1; JOINED.  
 DR EMBL; A9013902; BAA33388.1; JOINED.  
 DR HSSP; P25942; ICDF.  
 DR MGD; MGI:109587; Opg.  
 DR INTERPRO; IPR000488; -.  
 DR INTERPRO; IPR001368; -.  
 DR PFAM; PF00020; TNFR\_C6; 3.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 DR PROSITE; PS50050; TNFR\_NGFR\_2; 2.  
 DR PRODOM; PD000771; -; 1.  
 KW Glycoprotein; Repeat; Cytokine; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 401 OSTEOPROTEGERIN.  
 FT DOMAIN 23 201 4 X TNFR-CYS.  
 FT REPEAT 23 63 TNFR-CYS 1.  
 FT REPEAT 64 106 TNFR-CYS 2.  
 FT REPEAT 107 143 TNFR-CYS 3.  
 FT REPEAT 144 201 TNFR-CYS 4.  
 FT DOMAIN 306 365 DEATH DOMAIN.  
 FT DISULFID 41 54 BY SIMILARITY.  
 FT DISULFID 44 62 BY SIMILARITY.  
 FT DISULFID 65 80 BY SIMILARITY.  
 FT DISULFID 83 97 BY SIMILARITY.  
 FT DISULFID 87 105 BY SIMILARITY.  
 FT DISULFID 118 142 BY SIMILARITY.  
 FT DISULFID 145 160 BY SIMILARITY.  
 FT CARBOHYD 98 98 POTENTIAL.  
 FT CARBOHYD 165 165 POTENTIAL.  
 FT CARBOHYD 178 178 POTENTIAL.  
 FT CARBOHYD 289 289 POTENTIAL.  
 FT VARIANT 138 138  
 FT VARIANT 161 161  
 FT VARIANT 165 165  
 FT VARIANT 165 165  
 FT VARIANT 288 288  
 FT VARIANT 296 296  
 FT SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;

Query Match 39.5%; Score 888; DB 11; Length 401;  
 Best local Similarity 86.8%; Pred. No. 1.4e-69;  
 Matches 151; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 228 KETFPKYLHYDETSKQHCPCPTKYLKQHCAPCPDHYTDSWHTSDEC 287

Db 21 QETLPPKYLHYDETSKQHCPCPTKYLKQHCAPCPDHYTDSWHTSDEC 80  
 QY 288 LYCSPVCKELQYKQECNRTHNRVCECKEGRYLEIEFCLKHSRCPGPGVVGAGTTPRNT 347  
 Db 81 VYCSPVCKELQYKQECNRTHNRVCECKEGRYLEIEFCLKHSRCPGPGVVGAGTTPRNT 140  
 QY 348 VCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTOKGNATHDNCISGNSSTOK 401  
 Db 141 VCKRCPDGFSSNETSSKAPCRKHTNCSVFGLLLTOKGNATHDNCISGNSSTOK 194  
 RESULT 4  
 008727 PRELIMINARY; PRT; 401 AA.  
 AC 008727;  
 DT 01-JUL-1997 (TEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)  
 DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)  
 DE (OCIF)  
 GN TNFRSF11B OR OPG.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=INTESTINE;  
 RX MEDLINE=97262071; PubMed=9108485;  
 RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,  
 RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,  
 RA Derose M., Elliott R., Colombero A., Tan H.-L., Trall G., Sullivan J.,  
 RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,  
 RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,  
 RA Suggs S., Boyle W.J.;  
 RT "Osteoprotegerin: a novel secreted protein involved in the regulation  
 RT of bone density.";  
 RL Cell 89:309-319(1997).  
 CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES  
 CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY  
 CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO  
 CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN  
 CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.  
 CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY  
 CC SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.  
 DR EMBL; U94330; AAB53707.1; -.  
 DR HSSP; P25942; ICDF.  
 DR INTERPRO; IPR001368; -.  
 DR PFAM; PF00020; TNFR\_C6; 4.  
 DR PROSITE; PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE; PS50050; TNFR\_NGFR\_2; 2.  
 DR PRODOM; PD000771; -; 1.  
 KW Glycoprotein; Repeat; Cytokine; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 401 OSTEOPROTEGERIN.  
 FT DOMAIN 23 201 4 X TNFR-CYS.  
 FT REPEAT 23 63 TNFR-CYS 1.  
 FT REPEAT 64 106 TNFR-CYS 2.  
 FT REPEAT 107 143 TNFR-CYS 3.  
 FT REPEAT 144 201 TNFR-CYS 4.  
 FT DOMAIN 306 365 DEATH DOMAIN.  
 FT DISULFID 41 54 BY SIMILARITY.  
 FT DISULFID 44 62 BY SIMILARITY.  
 FT DISULFID 65 80 BY SIMILARITY.  
 FT DISULFID 83 97 BY SIMILARITY.  
 FT DISULFID 87 105 BY SIMILARITY.  
 FT DISULFID 118 142 BY SIMILARITY.  
 FT DISULFID 145 160 BY SIMILARITY.  
 FT CARBOHYD 98 98 POTENTIAL.  
 FT CARBOHYD 165 165 POTENTIAL.

```

FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
SQ SEQUENCE 401 AA; 46192 MW; PEC6A31F1D4E573A CRC64;

Query Match 39.0%; Score 875; DB 11; Length 401;
Best Local Similarity 85.5%; Pred. No. 1.8e-68;
Matches 146; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 228 KETPPKYLHVDETHSHLLCDKCPGTYLKHCHTAKWTKYCAPCPDHYHDSWHTSDEC 287
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
21 QETPPKYLHVDPETGROLCDKCAPGTYLKHCHTAKWTKYCAPCPDHYHDSWHTSDEC 80
QY 288 LYCSVPVKELQYVQECNRTNHRVCECKEGRYLEIFELCKHRSPPGPGVVOAGTPERN 347
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
81 VYCSVPVKELQYVQECNRTNHRVCECKEGRYLEIFELCKHRSPPGPGVVOAGTPERN 140
QY 348 VCKRCPDGFNSFNTSSKAPCRKHTNCVFGLLLTQKGNATHNDCSGNSESTQ 400
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
141 VCKRCPDGFNSFNTSSKAPCRKHTNCVFGLLLTQKGNATHNDCSGNSESTQ 193

RESULT 5
Q9RIA4 PRELIMINARY; PRT; 437 AA.
AC Q9RIA4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE GAMMA HEAVY CHAIN OF MAB7 (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -
DR HSSP; P01842; 7FAB.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; 1g; 4.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 36.7%; Score 824.5; DB 11; Length 437;
Best Local Similarity 62.9%; Pred. No. 5.2e-64;
Matches 141; Conservative 43; Mismatches 35; Indels 5; Gaps 2;

QY 7 CPP--CPAPELLGSPVLEFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYDGEV 64
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 CKPCICTVPEV---SSVFIEPPKPKDVLITLTPKVTVCVVVDISKDDPEVQFSWFVDVE 273
QY 65 VHNAKTPREQYNSTRVYVSVLTVLHODMLNGKEYCKYKSNKALPAPIETKISKARGQP 124
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
274 VHTAQTPREQYNSTRVYVSVLTVLHODMLNGKEYCKYKSNKALPAPIETKISKARGQP 333
QY 125 REPOVYTLPSRDLTKNOVSLCLVGFYPSDITAVWESNGQPNYKTPPVLDSDGS 184
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
334 KAPOVYTIPTPKQMAKDKVSLTKMTIDFPEDITVWQNGQPAENYKNTQPTIMDTDGS 393
QY 185 FFYLSKLTVDKSRMOQGNVFSVCSVHEALHNHYTKLSLSPGK 228
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
394 YFYVSKLVQKSNWEAGNTFTCSVLHSLHNHHTKLSHSPGK 437

RESULT 6

```

```

O95407 PRELIMINARY; PRT; 300 AA.
AC O95407;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN DCR3 OR TR6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA Goddard A.D., Botstein D., Ashkenazi A.;
RT "Genomic amplification of a decoy receptor for Fas ligand in lung and
RT colon cancer.";
RL Nature 396:699-703(1998).
RN [2]
SEQUENCE FROM N.A.
RC TISSUE-BLOOD;
RX Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT "A newly identified member of tumor necrosis factor receptor
RT superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL J. Biol. Chem. 274:13733-13736(1999).
[3]
SEQUENCE FROM N.A.
RC TISSUE-PANCREAS;
RX Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RX Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT "Overexpression of M68/Dcr3 in human gastrointestinal tract tumors
RT independent of gene amplification and its location in a four-gene
RT cluster.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR EMBL; AF104419; AAD03056.1; -.
DR EMBL; AF134240; AAD29688.1; -.
DR EMBL; AF217796; AAF35244.1; -.
DR EMBL; AF217793; AAF33685.1; -.
DR EMBL; AF217794; AAF33686.1; -.
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90ABE33718449AF CRC64;

Query Match 19.3%; Score 433; DB 4; Length 300;
Best Local Similarity 42.9%; Pred. No. 4.9e-30;
Matches 69; Conservative 29; Mismatches 63; Indels 0; Gaps 0;

QY 233 KYLHYDEETSHLLCDKCPGTYLKHCHTAKWTKYCAPCPDHYHDSWHTSDECYCSP 292
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 PTPWRDAETGERLUVCAQCPGTFVQRCRDSPTTCGCPRHRYTFWNLECRVCNV 93
QY 293 VKELQYVQECNRTNHRVCECKEGRYLEIFELCKHRSPPGPGVVOAGTPERNVCKRC 352
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
94 LGGEREEARACHATHNRACRCRTGFFAHAGFCLEHASCPPGAGVIAPGIFSQNTQCP 153
QY 353 PDGFFSNETSCKAPCRKHTNCVFGLLLTQKGNATHNDCS 393
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
154 PGGTFSSSSSEQCPHRNCTALGLALNVPGSSSHDTLCT 194

```







Best Local Similarity 31.4%; Pred. No. 3.3e-16;  
Matches 80; Conservative 33; Mismatches 110; Indels 32; Gaps 10;

QY 3 KHTHCP-----PCPAPELLGGPSVLFPPKPK-----DTLMISRTPEVT 41  
Db 119 KHTNPSQDVTPCPVPSTPTPTSPST-PTTPSPSCCHPRLSLHRLPALEDLLIGSSEANT 177

QY 42 CVVVDVSHEDPEYKFNWYDGVVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGKEYK 101  
Db 178 CTLTGL-ROASGVTFWTSSGR--SAVOGPPERDLGGCVSVSSVLPGCAEPNHHGKTFT 234

QY 102 KCVSNKALPAPIEKTISKAKGQPREQVYTLPPSRDELTKNQ-VSLTCLVKGYGPPSDIAV 160  
Db 235 CTAAYPESTPLTATLSKS-GNTFRPEVHLLPPSEELALNELVTLTCLARGSPKDVLV 293

QY 161 EWESNGQ--PENNYKTTPPVLD--SDG--SFFLYSKLTVDKSRWQGNVFCSCVYHHEALHN 215  
Db 294 RWLQSGQLPREKYLTVASRQSPSGTTFVAVTSILRVAEDWKKGDTFSCWVGHEALPL 353

QY 216 HYTKSLSPGKET 230  
Db 354 AFTQKTIIDRLAGKPT 368

RESULT 13  
QY0544 PRELIMINARY; PRT; 684 AA.

AC QY0544;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.  
OS Ginglymostoma cirratum (Nurse shark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;  
OC Ginglymostomatidae; Ginglymostoma.  
OX NCBI\_TaxID=7801;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RX MEDLINE=95183140; PubMed=7877689;  
RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,  
RA Flajnik M.F.;  
RT "A new antigen receptor gene family that undergoes rearrangement and  
RT extensive somatic diversification in sharks.";  
RL Nature 374:168-173(1995).  
DR EMBL: U18701; AAB48195.1; -.  
DR HSSP: P01857; 1FC1.  
DR INTERPRO: IPR003006; -.  
DR PFAM: PF00047; Ig; 6.  
DR PROSITE: PS00290; IG\_MHC; UNKNOWN\_3.  
KW Signal.  
FT SIGNAL. 1 18 POTENTIAL.  
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.  
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;

Query Match 12.1%; Score 272.5; DB 13; Length 684;  
Best Local Similarity 33.3%; Pred. No. 1.3e-15;  
Matches 71; Conservative 35; Mismatches 98; Indels 9; Gaps 7;

QY 20 SVFLFPKPKDITLMISRTPEVTCTVVDVSHEDPE-VKFNWYDGVVHNAKTPREEQYN 78  
Db 459 SVSLLKP-PREETWQTQATIVEIV---YSDLENIKVFWQVGVKRRKGVETQNPWSG 514

QY 79 STYRVSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKAK-GOPREQVYTLPPSRD 137  
Db 515 SKSTIVSKLVAMASEWDSGTEYVCLVEDSELPTPVKASIRKANVSQMHPPKVLLHPSTD 574

QY 138 EL-TKNQVSTCLVKGYGPPSDIAVESNGQ--PENNYKTTPPVLDSDGSFFLYSKLTVDK 195  
Db 575 EIDTENSATLMCLATNPHPAEIYVGMANDTLLDGSYRTQVDSEKSGSGSFVTDRLRLTA 634

QY 196 SRWQGNVFCSCVYHHEALHNHYTKSLSLSPGK 228  
Db 635 AEWNSTDTTSCVGVHPSL-NRDLIRSTNKSNGK 666

RESULT 14  
QYNNPP6 PRELIMINARY; PRT; 416 AA.

AC QYNNPP6;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Aufray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,  
RA Lehrach H., Poustka A., Lundeberg J.;  
RT "The European IMAGE consortium for integrated molecular analysis of  
RT human gene transcripts.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL389978; CAB97534.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8AB4687 CRC64;

Query Match 12.1%; Score 271; DB 4; Length 416;  
Best Local Similarity 32.8%; Pred. No. 1e-15;  
Matches 77; Conservative 34; Mismatches 112; Indels 12; Gaps 9;

QY 2 DKHTGCPCPAPELGGPSVLFPPKPKDTLMISRTPEVTCTVVDVSHEDPEVKFNWYVD 61  
Db 172 DVTVPCTVPPPPCC-HPRLSLHRLPALED-LLIGSEANLCTLGL-RDASGATFTWTPS 228

QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKAK 421  
Db 229 SGK--SAVOGPPERDLGGCVSVSSVLPGCAQPNWHGTFCTAAHPELKTPLTANITKS- 285

QY 122 GQPREQVYTLPPSRDELTKNQ-VSLTCLVKGYGPPSDIAVESNGQ--PENNYKTTTPV 178  
Db 286 GNTFRPEVHLLPPSEELALNELVTLTCLARGSPKDVLRWLQSGQELPREKYLTVASR 345

QY 179 LD-SDG--SFFLYSKLTVDKSRWQGNVFCSCVYHHEALHNHYTKSLSLSPGKET 230  
Db 346 QEPFSGTTFVAVTSILRVAEDWKKGDTFSCWVGHEALPLAFTQKTIIDRLAGKPT 400

RESULT 15  
QY606 PRELIMINARY; PRT; 616 AA.

AC QY606;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE RECEPTOR ACTIVATOR OF NF-KAPPA-B PRECURSOR (TNF-RELATED ACTIVATION-  
DE INDUCED CYTOKINE RECEPTOR) (RANK).  
GN TNFRSF1A OR RANK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98032977; PubMed=9367155;  
RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,  
RA Tometsko M.E., Roux E.R., Teepe M.C., DuBoise R.F., Cosman D.,



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:48 ; Search time 40.97 Seconds  
(without alignments)  
316.083 Million cell updates/sec

Title: US-09-389-782a-8

Perfect score: 2246

Sequence: 1 MDKTHTCPPCAPPELLGPGS.....CKGNATHDNCNSGSESTQK 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query %	Length	DB ID	Description
1	1233	54.9	330	1	GCL_HUMAN	P01857 homo sapien
2	1142.5	50.9	326	1	GC2_HUMAN	P01859 homo sapien
3	1135	50.5	327	1	GC4_HUMAN	P01861 homo sapien
4	1126	50.1	290	1	GC3_HUMAN	P01860 homo sapien
5	918.5	40.9	323	1	GC_RABIT	P01870 oryctolagus
6	889	39.6	329	1	GC2_CAVPO	P01862 cavia porce
7	845.5	37.6	329	1	GC3_MOUSE	P22436 mus musculus
8	838	37.3	333	1	GC3_RAT	P20761 rattus norv
9	835.5	37.2	398	1	GC3M_MOUSE	P03987 mus musculus
10	818.5	36.4	326	1	GCL_RAT	P20759 rattus norv
11	817.5	36.4	324	1	GCL_MOUSE	P01868 mus musculus
12	812.5	36.2	393	1	GC1M_MOUSE	P01869 mus musculus
13	809.5	36.0	329	1	GCC_RAT	P20762 rattus norv
14	809	36.0	330	1	GCAA_MOUSE	P01863 mus musculus
15	804	35.8	399	1	GCAM_MOUSE	P01865 mus musculus
16	802	35.7	335	1	GCAB_MOUSE	P01864 mus musculus
17	785.5	35.0	322	1	GCA_RAT	P20760 rattus norv
18	779	34.7	336	1	GCB_MOUSE	P01866 mus musculus
19	774	34.5	405	1	GCEB_MOUSE	P01867 mus musculus
20	362	16.1	454	1	MUC_HUMAN	P01871 homo sapien
21	361	16.1	455	1	MUC_MOUSE	P01872 mus musculus
22	360	16.0	391	1	MUCB_HUMAN	P04220 homo sapien
23	359	16.0	421	1	EPC_MOUSE	P06336 mus musculus
24	358	15.9	429	1	EPC_RAT	P01855 rattus norv
25	352	15.7	428	1	EPC_HUMAN	P01854 homo sapien
26	347	15.4	476	1	MUCM_MOUSE	P01873 mus musculus
27	346	15.4	458	1	MUC_RABIT	P03988 oryctolagus
28	336	15.0	450	1	MUC_CANFA	P01874 canis famil
29	336	15.0	454	1	MUC_MESAU	P06337 mesocricetu
30	333	14.8	457	1	MUC_SUNMU	P20768 sunus muri
31	333	14.8	461	1	TNR2_HUMAN	P20333 homo sapien
32	332.5	14.8	474	1	TNR2_MOUSE	P25119 mus musculus
33	332	14.8	479	1	MUCM_RABIT	P04221 oryctolagus

## RESULT 1

ID	GCL_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG GAMMA-1 CHAIN C REGION.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RL	Biochemistry 9:3161-3170(1970).			
RN	[3]			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RL	Biochemistry 9:3171-3181(1970).			
RN	[4]			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The chymotryptic peptides of the H-chain, alignment of the tryptic			
RT	peptides and discussion of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RN	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[6]			
RP	DISULFIDE BONDS.			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X. Intrachain disulfide bonds."			

P23085 heterodontu  
P01879 oryctolagus  
P23087 heterodontu  
P01875 gallus gall  
P27312 mus musculu  
P23086 heterodontu  
P23088 heterodontu  
P23084 heterodontu  
P01876 homo sapien  
P20758 gorilla gor  
P01877 homo sapien  
P25942 homo sapien

## ALIGNMENTS

34 303 13.5 438 1 HVC2\_HETFR  
35 299.5 13.3 299 1 ALC\_RABIT  
36 293.5 13.1 438 1 HVCS\_HETFR  
37 287.5 12.8 446 1 MUC\_CHICK  
38 284 12.6 289 1 CD40\_MOUSE  
39 282.5 12.6 393 1 HVC3\_HETFR  
40 278.5 12.4 461 1 HVCM\_HETFR  
41 278 12.4 370 1 HVC1\_HETFR  
42 276 12.3 353 1 ALC1\_HUMAN  
43 271.5 12.1 353 1 ALC1\_GORGO  
44 271 12.1 340 1 ALC2\_HUMAN  
45 264 11.8 277 1 CD40\_HUMAN

Biochemistry 9:3188-3196(1970).  
 [7]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=77070267; PubMed=1002129;  
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 RT "Rule of antibody structure. The primary structure of a monoclonal  
 RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and  
 RT characterization of the protein, the L- and H-chains, the  
 RT cyanogen bromide cleavage products, and the disulfide bridges.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RX MEDLINE=81208100; PubMed=7236608;  
 RA Deisenhofer J.;  
 RT "Crystallographic refinement and atomic models of a human Fc fragment  
 RT and its complex with fragment B of protein A from Staphylococcus  
 RT aureus at 2.9- and 2.8-A resolution.";  
 RL Biochemistry 20:2361-2370(1981).  
 CC -!- MISCELLANEOUS: NIE HAS THE G1M(17) ALLOTYPIC MARKER, 97-K, & THE  
 CC G1M(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE G1M(3)  
 CC MARKER & THE G1M (NON-1) MARKERS.  
 CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC 35,116,198,269 & 272.  
 CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
 CC 268-272.  
 CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC RESIDUES 198,267&272.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; J00228; AAC82527.1; ALT\_INIT.  
 DR PIR; A02146; GHU.  
 DR PDB; 1FC1; 15-JUL-92.  
 DR PDB; 1FC2; 15-JUL-92.  
 DR MIN; 147100; -.  
 DR INTERPRO; IPR000495; -.  
 DR INTERPRO; IPR003006; -.  
 DR PFAM; PF00047; Ig; 3.  
 DR PROSITE; PS00290; IG\_MHC; 2.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 3D-structure.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 98 CH1.  
 FT HINGE.  
 FT DOMAIN 99 110 CH2.  
 FT DOMAIN 111 223 CH2.  
 FT DOMAIN 224 330 CH3.  
 FT DISULFID 27 83  
 FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).  
 FT DISULFID 144 204  
 FT DISULFID 250 308  
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .).  
 FT VARIANT 97 97 K -> R (IN G1M(3) MARKER).  
 FT D -> E (IN G1M(NON-1) MARKER).  
 FT /FTId=VAR\_003886.  
 FT /FTId=VAR\_003887.  
 FT L -> M (IN G1M(NON-1) MARKER).  
 FT /FTId=VAR\_003888.  
 FT REMOVED POST-TRANSLATIONALLY.  
 FT VARIANT 239 239  
 FT VARIANT 241 241  
 FT MOD\_RES 330 330  
 FT STRAND 123 126  
 FT HELIX 130 134  
 FT TURN 136 137  
 FT STRAND 141 148  
 FT STRAND 158 162  
 FT TURN 163 164

FT STRAND 165 166  
 FT STRAND 175 178  
 FT STRAND 183 190  
 FT HELIX 193 197  
 FT TURN 198 199  
 FT STRAND 202 206  
 FT STRAND 215 219  
 FT STRAND 227 227  
 FT STRAND 230 234.  
 FT STRAND 238 240  
 FT TURN 241 242  
 FT STRAND 245 256  
 FT STRAND 260 266  
 FT TURN 267 268  
 FT STRAND 269 270  
 FT STRAND 274 276  
 FT STRAND 280 281  
 FT TURN 283 284  
 FT STRAND 287 296  
 FT STRAND 297 301  
 FT HELIX 302 303  
 FT TURN 306 312  
 FT TURN 313 314  
 FT TURN 316 317  
 FT STRAND 320 324  
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;  
 Query Match 54.9%; Score 1233; DB 1; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-78;  
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 DKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYVD 61  
 DB 104 DKTHCTPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYVD 163  
 QY 62 GVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121  
 DB 164 GVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 223  
 QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181  
 DB 224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 283  
 QY 182 DGSFELYSLKLVDSKRWQOGNVSFSCSVNHEALHNHYTQKSLSLSPGK 228  
 DB 284 DGSFELYSLKLVDSKRWQOGNVSFSCSVNHEALHNHYTQKSLSLSPGK 330  
 RESULT 2  
 GC2\_HUMAN  
 ID GC2\_HUMAN STANDARD; PRT; 326 AA.  
 AC P01859;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-2 CHAIN C REGION.  
 GN IGHG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82197621; PubMed=6804948;  
 RA Ellison J.W., Hood L.E.;  
 RT "Linkage and sequence homology of two human immunoglobulin gamma  
 RT heavy chain constant region genes.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).  
 RN [2]  
 RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).  
 RX MEDLINE=81007873; PubMed=6774012;  
 RA Wang A.-C., Tung E., Fudenberg H.H.;  
 RT "The primary structure of a human IgG2 heavy chain: genetic,

evolutionary, and functional implications.";

[3] J. Immunol. 125:1048-1054(1980).

RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).

RX MEDLINE=80001357; PubMed=113060;

RA Connell G.E., Parr D.M., Hofmann T.;

RT "The amino acid sequences of the three heavy chain constant region domains of a human IgG2 myeloma protein.";

RL Can. J. Biochem. 57:758-767(1979).

RN [4].

RP SEQUENCE OF 238-275 (ZIE).

RX MEDLINE=80114419; PubMed=118920;

RA Hofmann T., Parr D.M.;

RT "A note of the amino acid sequence of residues 381-391 of human immunoglobulins gamma chains.";

RL Mol. Immunol. 16:923-925(1979).

RN [5].

RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).

RX Hofmann T., Parr D.M.;

RA Submitted (MAR-1980) to the PIR data bank.

RN [6].

RP SEQUENCE OF 1-121 (DOT).

RX MEDLINE=95255298; PubMed=7737190;

RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;

RT "Characterization of the two unique human anti-flavin monoclonal immunoglobulins.";

RL Eur. J. Biochem. 228:886-893(1995).

RN [7].

RP DISULFIDE BONDS.

RX MEDLINE=72033500; PubMed=4940472;

RA Milstein C., Frangione B.;

RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";

RL Biochem. J. 121:217-225(1971).

RN [8].

RP DISULFIDE BONDS.

RX MEDLINE=69064124; PubMed=5782707;

RA Frangione B., Milstein C., Pink J.R.L.;

RT "Structure studies of immunoglobulin G.";

RL Nature 221:145-148(1969).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; V00554; CAA23814.1; -

DR EMBL; V00554; CAA23815.1; -

DR EMBL; V00554; CAA23816.1; -

DR EMBL; V00554; CAA23817.1; -

DR PIR; A02148; G2HU.

DR MIM; 147110; -

DR INTERPRO; IPR000495; -

DR INTERPRO; IPR003006; -

DR PFAM; PF00047; ig; 3.

DR PROSITE; PS00290; IG\_MHC; 2.

KW Immunoglobulin domain; Immunoglobulin C region.

FT NON\_TER 1 1

FT DOMAIN 1 98 CHI.

FT DOMAIN 99 110 HINGE.

FT DOMAIN 111 219 CH2.

FT DOMAIN 220 326 CH3.

FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).

FT DISULFID 27 83

FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 140 200

FT DISULFID 246 304

FT VARIANT 60

S -> A (IN MYELOMA PROTEINS TIL & ZIE).

FT SITE 156 156 /FTid=VAR\_003889.

FT MOD\_RES 326 326 AT OR NEAR THE COMPLEMENT-BINDING SITE.

SQ SEQUENCE 326 AA; 35884 MW; 8310876C6878CF9C CRC64;

Best Local Similarity 94.1%; Pred. No. 2.1e-72;

Query Match 50.9%; Score 1142.5; DB 1; Length 326;

Matches 209; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

QY 7 CPCCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 56

Db 106 CPCCPAAPP-VAGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVQFNWYVDGVEVH 64

QY 67 NAKTKPREEOYNSTYRVVSVLTVLDHDLANKKEYCKVSKNKPAPLPIETKISKAKGPPE 26

Db 165 NAKTKPREEQFNSTFRVSVLTVVHDLANKKEYCKVSKNKPAPLPIETKISKAKGPPE 24

QY 127 PQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPVLDSDGSF 186

Db 225 PQVYTLPPSRDEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPVLDSDGSF 284

QY 187 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQKSLSLSPGK 228

Db 285 LYSKLTVDKSRWQGNVFCFSVMHEALHNHYTQKSLSLSPGK 326

RESULT 3

GC4\_HUMAN STANDARD; PRT; 327 AA.

ID AC P01861;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE IG GAMMA-4 CHAIN C REGION.

GN IGHG4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83157104; PubMed=6299662;

RA Ellison J.W., Buxbaum J.N., Hood L.E.;

RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";

RL DNA 1:11-18(1981).

RN [2]

RP SEQUENCE OF 1-30 AND 81-326.

RX MEDLINE=70207560; PubMed=4192699;

RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;

RT "Human immunoglobulin subclasses. Partial amino acid sequence of the constant region of a gamma 4 chain.";

RL Biochem. J. 117:33-47(1970).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; K01316; AAB59394.1; ALT\_INIT.

DR PIR; A02150; G4HU.

DR MIM; 147130; -

DR INTERPRO; IPR000495; -

DR INTERPRO; IPR003006; -

DR PFAM; PF00047; ig; 3.

DR PROSITE; PS00290; IG\_MHC; 2.

KW Immunoglobulin domain; Immunoglobulin C region.

FT NON\_TER 1 1

FT DOMAIN 1 98 CHI.

FT DOMAIN 99 110 HINGE.

FT DOMAIN 111 220 CH2.

FT DOMAIN 221 327 CH3.  
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 83  
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 141 201  
FT DISULFID 247 305  
SQ SEQUENCE 327 AA; 35940 MW; 3EDB0811EF208E7A CRC64;  
  
Query Match 50.5%; Score 1135; DB 1; Length 327;  
Best Local Similarity 93.7%; Pred. No. 7.1e-72;  
Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;  
  
QY 7 CPCCPAPELLGSPVFLPPKPKDTLMSRTPVTCVVDVSHDEPKVFNWYDGVGVH 66  
DB 106 CPSCPAPELGLSPVFLPPKPKDTLMSRTPVTCVVDVSHDEPKVFNWYDGVGVH 165  
  
QY 67 NAKTKPREEQNSTYRVVSVLTVLHODWLNKGYKCKVSNKALPAPIETKTSKAGQPRE 126  
DB 166 NAKTKPREEQNSTYRVVSVLTVLHODWLNKGYKCKVSNKALPAPIETKTSKAGQPRE 225  
  
QY 127 PQVYTLPSRDELTKNOVSLCLVKGFPYSDIAVWESNGQPNKYKTPPVLDSDGSFF 186  
DB 226 PQVYTLPSRDELTKNOVSLCLVKGFPYSDIAVWESNGQPNKYKTPPVLDSDGSFF 285  
  
QY 187 LYSKLTVDKSRWQGNVFCSCVMHHEALNHYTKLSLSPGK 228  
DB 286 LYSRLTVDKSRWQGNVFCSCVMHHEALNHYTKLSLSPGK 327

RESULT 4  
GC3\_HUMAN STANDARD; PRT; 290 AA.  
AC P01860;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE IG GAMMA-3 CHAIN C REGION (HEAVY CHAIN DISEASE PROTEIN) (HDC).  
GN IGHG3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
[1]  
RX MEDLINE-81021548; PubMed-6774747;  
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;  
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:  
RT gamma 3 heavy-chain disease protein wis.";  
RL Biochemistry 19:4304-4308(1980).  
[2]  
RN REVISIONS TO 12-97 OF PROTEIN WIS.  
RX MEDLINE-77118561; PubMed-402363;  
RA Michaelson T.E., Frangione B., Franklin E.C.;  
RT "Primary structure of the 'hinge' region of human IgG3. Probable  
RT quadruplication of a 13-amino acid residue basic unit.";  
RL J. Biol. Chem. 252:883-889(1977).  
[3]  
RN REVISIONS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).  
RX MEDLINE-77021516; PubMed-823945;  
RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;  
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.  
RT Structure of the Fc fragment of immunoglobulin G3";  
RL Biochem. Biophys. Res. Commun. 71:907-914(1976).  
[4]  
RN SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).  
RX MEDLINE-82247835; PubMed-6808505;  
RA Alexander A., Steinmetz M., Barritault D., Frangione B.,  
RA Franklin E.C., Hood L., Buxbaum J.N.;  
RT "Gamma Heavy chain disease in man: cDNA sequence supports partial  
RT gene deletion model";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).  
CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA

CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11  
CC NORMALLY PRESENT IN THE HINGE REGION.  
CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.  
CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE  
CC REF.2.  
CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION  
CC AND ALL OF THE CH1 REGION.  
CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL  
CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL  
CC GAMMA-3 HEAVY CHAINS.  
CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM  
CC OR ANOTHER GAMMA CHAIN SUBCLASS.  
CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR  
CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE  
CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE  
CC SEGMENT (12-28).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
CC -----  
DR EMBL: J00231; AAA52805.1; ALT\_SEQ.  
DR PIR: A02149; G3HUW1.  
DR MIM: 147120;  
DR INTERPRO: IPR000495;  
DR INTERPRO: IPR003006;  
DR PFAM: PF00047; Ig; 2.  
DR PROSITE: PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
FT DOMAIN 12 73  
FT DOMAIN 74 183  
FT DOMAIN 184 289  
FT REPEAT 29 43  
FT REPEAT 44 58  
FT REPEAT 59 73  
FT MOD\_RES 1 1  
FT CARBOHYD 6 6  
FT DISULFID 7 7  
FT DISULFID 24 24  
FT DISULFID 27 27  
FT DISULFID 33 33  
FT DISULFID 39 39  
FT DISULFID 42 42  
FT DISULFID 48 48  
FT DISULFID 54 54  
FT DISULFID 57 57  
FT DISULFID 63 63  
FT DISULFID 69 69  
FT DISULFID 72 72  
FT CARBOHYD 140 140  
FT MOD\_RES 290 290  
FT VARIANT 126 127  
FT VARIANT 134 134  
FT VARIANT 139 139  
FT VARIANT 182 182  
FT VARIANT 227 227  
FT VARIANT 227 227  
FT VARIANT 279 279  
FT SEQUENCE 290 AA; 32331 MW; E69CBC95705B2F46 CRC64;  
  
P -> L (IN OMM).  
F -> Y (IN OMM).  
T -> A (IN OMM).  
S -> N (IN OMM).  
MISSING (IN ZUC).  
F -> Y (IN OMM).  
/FtId=VAR\_003890.  
/FtId=VAR\_003891.  
/FtId=VAR\_003892.  
/FtId=VAR\_003893.  
/FtId=VAR\_003894.  
/FtId=VAR\_003895.  
/FtId=VAR\_003896.  
/FtId=VAR\_003897.  
/FtId=VAR\_003898.  
/FtId=VAR\_003899.  
/FtId=VAR\_003900.  
/FtId=VAR\_003901.  
/FtId=VAR\_003902.  
/FtId=VAR\_003903.  
/FtId=VAR\_003904.  
/FtId=VAR\_003905.  
/FtId=VAR\_003906.  
/FtId=VAR\_003907.  
/FtId=VAR\_003908.  
/FtId=VAR\_003909.  
/FtId=VAR\_003910.  
/FtId=VAR\_003911.  
/FtId=VAR\_003912.  
/FtId=VAR\_003913.  
/FtId=VAR\_003914.  
/FtId=VAR\_003915.  
/FtId=VAR\_003916.  
/FtId=VAR\_003917.  
/FtId=VAR\_003918.  
/FtId=VAR\_003919.  
/FtId=VAR\_003920.  
/FtId=VAR\_003921.  
/FtId=VAR\_003922.  
/FtId=VAR\_003923.  
/FtId=VAR\_003924.  
/FtId=VAR\_003925.  
/FtId=VAR\_003926.  
/FtId=VAR\_003927.  
/FtId=VAR\_003928.  
/FtId=VAR\_003929.  
/FtId=VAR\_003930.  
/FtId=VAR\_003931.  
/FtId=VAR\_003932.  
/FtId=VAR\_003933.  
/FtId=VAR\_003934.  
/FtId=VAR\_003935.  
/FtId=VAR\_003936.  
/FtId=VAR\_003937.  
/FtId=VAR\_003938.  
/FtId=VAR\_003939.  
/FtId=VAR\_003940.  
/FtId=VAR\_003941.  
/FtId=VAR\_003942.  
/FtId=VAR\_003943.  
/FtId=VAR\_003944.  
/FtId=VAR\_003945.  
/FtId=VAR\_003946.  
/FtId=VAR\_003947.  
/FtId=VAR\_003948.  
/FtId=VAR\_003949.  
/FtId=VAR\_003950.  
/FtId=VAR\_003951.  
/FtId=VAR\_003952.  
/FtId=VAR\_003953.  
/FtId=VAR\_003954.  
/FtId=VAR\_003955.  
/FtId=VAR\_003956.  
/FtId=VAR\_003957.  
/FtId=VAR\_003958.  
/FtId=VAR\_003959.  
/FtId=VAR\_003960.  
/FtId=VAR\_003961.  
/FtId=VAR\_003962.  
/FtId=VAR\_003963.  
/FtId=VAR\_003964.  
/FtId=VAR\_003965.  
/FtId=VAR\_003966.  
/FtId=VAR\_003967.  
/FtId=VAR\_003968.  
/FtId=VAR\_003969.  
/FtId=VAR\_003970.  
/FtId=VAR\_003971.  
/FtId=VAR\_003972.  
/FtId=VAR\_003973.  
/FtId=VAR\_003974.  
/FtId=VAR\_003975.  
/FtId=VAR\_003976.  
/FtId=VAR\_003977.  
/FtId=VAR\_003978.  
/FtId=VAR\_003979.  
/FtId=VAR\_003980.  
/FtId=VAR\_003981.  
/FtId=VAR\_003982.  
/FtId=VAR\_003983.  
/FtId=VAR\_003984.  
/FtId=VAR\_003985.  
/FtId=VAR\_003986.  
/FtId=VAR\_003987.  
/FtId=VAR\_003988.  
/FtId=VAR\_003989.  
/FtId=VAR\_003990.  
/FtId=VAR\_003991.  
/FtId=VAR\_003992.  
/FtId=VAR\_003993.  
/FtId=VAR\_003994.  
/FtId=VAR\_003995.  
/FtId=VAR\_003996.  
/FtId=VAR\_003997.  
/FtId=VAR\_003998.  
/FtId=VAR\_003999.  
/FtId=VAR\_004000.  
/FtId=VAR\_004001.  
/FtId=VAR\_004002.  
/FtId=VAR\_004003.  
/FtId=VAR\_004004.  
/FtId=VAR\_004005.  
/FtId=VAR\_004006.  
/FtId=VAR\_004007.  
/FtId=VAR\_004008.  
/FtId=VAR\_004009.  
/FtId=VAR\_004010.  
/FtId=VAR\_004011.  
/FtId=VAR\_004012.  
/FtId=VAR\_004013.  
/FtId=VAR\_004014.  
/FtId=VAR\_004015.  
/FtId=VAR\_004016.  
/FtId=VAR\_004017.  
/FtId=VAR\_004018.  
/FtId=VAR\_004019.  
/FtId=VAR\_004020.  
/FtId=VAR\_004021.  
/FtId=VAR\_004022.  
/FtId=VAR\_004023.  
/FtId=VAR\_004024.  
/FtId=VAR\_004025.  
/FtId=VAR\_004026.  
/FtId=VAR\_004027.  
/FtId=VAR\_004028.  
/FtId=VAR\_004029.  
/FtId=VAR\_004030.  
/FtId=VAR\_004031.  
/FtId=VAR\_004032.  
/FtId=VAR\_004033.  
/FtId=VAR\_004034.  
/FtId=VAR\_004035.  
/FtId=VAR\_004036.  
/FtId=VAR\_004037.  
/FtId=VAR\_004038.  
/FtId=VAR\_004039.  
/FtId=VAR\_004040.  
/FtId=VAR\_004041.  
/FtId=VAR\_004042.  
/FtId=VAR\_004043.  
/FtId=VAR\_004044.  
/FtId=VAR\_004045.  
/FtId=VAR\_004046.  
/FtId=VAR\_004047.  
/FtId=VAR\_004048.  
/FtId=VAR\_004049.  
/FtId=VAR\_004050.  
/FtId=VAR\_004051.  
/FtId=VAR\_004052.  
/FtId=VAR\_004053.  
/FtId=VAR\_004054.  
/FtId=VAR\_004055.  
/FtId=VAR\_004056.  
/FtId=VAR\_004057.  
/FtId=VAR\_004058.  
/FtId=VAR\_004059.  
/FtId=VAR\_004060.  
/FtId=VAR\_004061.  
/FtId=VAR\_004062.  
/FtId=VAR\_004063.  
/FtId=VAR\_004064.  
/FtId=VAR\_004065.  
/FtId=VAR\_004066.  
/FtId=VAR\_004067.  
/FtId=VAR\_004068.  
/FtId=VAR\_004069.  
/FtId=VAR\_004070.  
/FtId=VAR\_004071.  
/FtId=VAR\_004072.  
/FtId=VAR\_004073.  
/FtId=VAR\_004074.  
/FtId=VAR\_004075.  
/FtId=VAR\_004076.  
/FtId=VAR\_004077.  
/FtId=VAR\_004078.  
/FtId=VAR\_004079.  
/FtId=VAR\_004080.  
/FtId=VAR\_004081.  
/FtId=VAR\_004082.  
/FtId=VAR\_004083.  
/FtId=VAR\_004084.  
/FtId=VAR\_004085.  
/FtId=VAR\_004086.  
/FtId=VAR\_004087.  
/FtId=VAR\_004088.  
/FtId=VAR\_004089.  
/FtId=VAR\_004090.  
/FtId=VAR\_004091.  
/FtId=VAR\_004092.  
/FtId=VAR\_004093.  
/FtId=VAR\_004094.  
/FtId=VAR\_004095.  
/FtId=VAR\_004096.  
/FtId=VAR\_004097.  
/FtId=VAR\_004098.  
/FtId=VAR\_004099.  
/FtId=VAR\_004100.  
/FtId=VAR\_004101.  
/FtId=VAR\_004102.  
/FtId=VAR\_004103.  
/FtId=VAR\_004104.  
/FtId=VAR\_004105.  
/FtId=VAR\_004106.  
/FtId=VAR\_004107.  
/FtId=VAR\_004108.  
/FtId=VAR\_004109.  
/FtId=VAR\_004110.  
/FtId=VAR\_004111.  
/FtId=VAR\_004112.  
/FtId=VAR\_004113.  
/FtId=VAR\_004114.  
/FtId=VAR\_004115.  
/FtId=VAR\_004116.  
/FtId=VAR\_004117.  
/FtId=VAR\_004118.  
/FtId=VAR\_004119.  
/FtId=VAR\_004120.  
/FtId=VAR\_004121.  
/FtId=VAR\_004122.  
/FtId=VAR\_004123.  
/FtId=VAR\_004124.  
/FtId=VAR\_004125.  
/FtId=VAR\_004126.  
/FtId=VAR\_004127.  
/FtId=VAR\_004128.  
/FtId=VAR\_004129.  
/FtId=VAR\_004130.  
/FtId=VAR\_004131.  
/FtId=VAR\_004132.  
/FtId=VAR\_004133.  
/FtId=VAR\_004134.  
/FtId=VAR\_004135.  
/FtId=VAR\_004136.  
/FtId=VAR\_004137.  
/FtId=VAR\_004138.  
/FtId=VAR\_004139.  
/FtId=VAR\_004140.  
/FtId=VAR\_004141.  
/FtId=VAR\_004142.  
/FtId=VAR\_004143.  
/FtId=VAR\_004144.  
/FtId=VAR\_004145.  
/FtId=VAR\_004146.  
/FtId=VAR\_004147.  
/FtId=VAR\_004148.  
/FtId=VAR\_004149.  
/FtId=VAR\_004150.  
/FtId=VAR\_004151.  
/FtId=VAR\_004152.  
/FtId=VAR\_004153.  
/FtId=VAR\_004154.  
/FtId=VAR\_004155.  
/FtId=VAR\_004156.  
/FtId=VAR\_004157.  
/FtId=VAR\_004158.  
/FtId=VAR\_004159.  
/FtId=VAR\_004160.  
/FtId=VAR\_004161.  
/FtId=VAR\_004162.  
/FtId=VAR\_004163.  
/FtId=VAR\_004164.  
/FtId=VAR\_004165.  
/FtId=VAR\_004166.  
/FtId=VAR\_004167.  
/FtId=VAR\_004168.  
/FtId=VAR\_004169.  
/FtId=VAR\_004170.  
/FtId=VAR\_004171.  
/FtId=VAR\_004172.  
/FtId=VAR\_004173.  
/FtId=VAR\_004174.  
/FtId=VAR\_004175.  
/FtId=VAR\_004176.  
/FtId=VAR\_004177.  
/FtId=VAR\_004178.  
/FtId=VAR\_004179.  
/FtId=VAR\_004180.  
/FtId=VAR\_004181.  
/FtId=VAR\_004182.  
/FtId=VAR\_004183.  
/FtId=VAR\_004184.  
/FtId=VAR\_004185.  
/FtId=VAR\_004186.  
/FtId=VAR\_004187.  
/FtId=VAR\_004188.  
/FtId=VAR\_004189.  
/FtId=VAR\_004190.  
/FtId=VAR\_004191.  
/FtId=VAR\_004192.  
/FtId=VAR\_004193.  
/FtId=VAR\_004194.  
/FtId=VAR\_004195.  
/FtId=VAR\_004196.  
/FtId=VAR\_004197.  
/FtId=VAR\_004198.  
/FtId=VAR\_004199.  
/FtId=VAR\_004200.  
/FtId=VAR\_004201.  
/FtId=VAR\_004202.  
/FtId=VAR\_004203.  
/FtId=VAR\_004204.  
/FtId=VAR\_004205.  
/FtId=VAR\_004206.  
/FtId=VAR\_004207.  
/FtId=VAR\_004208.  
/FtId=VAR\_004209.  
/FtId=VAR\_004210.  
/FtId=VAR\_004211.  
/FtId=VAR\_004212.  
/FtId=VAR\_004213.  
/FtId=VAR\_004214.  
/FtId=VAR\_004215.  
/FtId=VAR\_004216.  
/FtId=VAR\_004217.  
/FtId=VAR\_004218.  
/FtId=VAR\_004219.  
/FtId=VAR\_004220.  
/FtId=VAR\_004221.  
/FtId=VAR\_004222.  
/FtId=VAR\_004223.  
/FtId=VAR\_004224.  
/FtId=VAR\_004225.  
/FtId=VAR\_004226.  
/FtId=VAR\_004227.  
/FtId=VAR\_004228.  
/FtId=VAR\_004229.  
/FtId=VAR\_004230.  
/FtId=VAR\_004231.  
/FtId=VAR\_004232.  
/FtId=VAR\_004233.  
/FtId=VAR\_004234.  
/FtId=VAR\_004235.  
/FtId=VAR\_004236.  
/FtId=VAR\_004237.  
/FtId=VAR\_004238.  
/FtId=VAR\_004239.  
/FtId=VAR\_004240.  
/FtId=VAR\_004241.  
/FtId=VAR\_004242.  
/FtId=VAR\_004243.  
/FtId=VAR\_004244.  
/FtId=VAR\_004245.  
/FtId=VAR\_004246.  
/FtId=VAR\_004247.  
/FtId=VAR\_004248.  
/FtId=VAR\_004249.  
/FtId=VAR\_004250.  
/FtId=VAR\_004251.  
/FtId=VAR\_004252.  
/FtId=VAR\_004253.  
/FtId=VAR\_004254.  
/FtId=VAR\_004255.  
/FtId=VAR\_004256.  
/FtId=VAR\_004257.  
/FtId=VAR\_004258.  
/FtId=VAR\_004259.  
/FtId=VAR\_004260.  
/FtId=VAR\_004261.  
/FtId=VAR\_004262.  
/FtId=VAR\_004263.  
/FtId=VAR\_004264.  
/FtId=VAR\_004265.  
/FtId=VAR\_004266.  
/FtId=VAR\_004267.  
/FtId=VAR\_004268.  
/FtId=VAR\_004269.  
/FtId=VAR\_004270.  
/FtId=VAR\_004271.  
/FtId=VAR\_004272.  
/FtId=VAR\_004273.  
/FtId=VAR\_004274.  
/FtId=VAR\_004275.  
/FtId=VAR\_004276.  
/FtId=VAR\_004277.  
/FtId=VAR\_004278.  
/FtId=VAR\_004279.  
/FtId=VAR\_004280.  
/FtId=VAR\_004281.  
/FtId=VAR\_004282.  
/FtId=VAR\_004283.  
/FtId=VAR\_004284.  
/FtId=VAR\_004285.  
/FtId=VAR\_004286.  
/FtId=VAR\_004287.  
/FtId=VAR\_004288.  
/FtId=VAR\_004289.  
/FtId=VAR\_004290.  
/FtId=VAR\_004291.  
/FtId=VAR\_004292.  
/FtId=VAR\_004293.  
/FtId=VAR\_004294.  
/FtId=VAR\_004295.  
/FtId=VAR\_004296.  
/FtId=VAR\_004297.  
/FtId=VAR\_004298.  
/FtId=VAR\_004299.  
/FtId=VAR\_004300.  
/FtId=VAR\_004301.  
/FtId=VAR\_004302.  
/FtId=VAR\_004303.  
/FtId=VAR\_004304.  
/FtId=VAR\_004305.  
/FtId=VAR\_004306.  
/FtId=VAR\_004307.  
/FtId=VAR\_004308.  
/FtId=VAR\_004309.  
/FtId=VAR\_004310.  
/FtId=VAR\_004311.  
/FtId=VAR\_004312.  
/FtId=VAR\_004313.  
/FtId=VAR\_004314.  
/FtId=VAR\_004315.  
/FtId=VAR\_004316.  
/FtId=VAR\_004317.  
/FtId=VAR\_004318.  
/FtId=VAR\_004319.  
/FtId=VAR\_004320.  
/FtId=VAR\_004321.  
/FtId=VAR\_004322.  
/FtId=VAR\_004323.  
/FtId=VAR\_004324.  
/FtId=VAR\_004325.  
/FtId=VAR\_004326.  
/FtId=VAR\_004327.  
/FtId=VAR\_004328.  
/FtId=VAR\_004329.  
/FtId=VAR\_004330.  
/FtId=VAR\_004331.  
/FtId=VAR\_004332.  
/FtId=VAR\_004333.  
/FtId=VAR\_004334.  
/FtId=VAR\_004335.  
/FtId=VAR\_004336.  
/FtId=VAR\_004337.  
/FtId=VAR\_004338.  
/FtId=VAR\_004339.  
/FtId=VAR\_004340.  
/FtId=VAR\_004341.  
/FtId=VAR\_004342.  
/FtId=VAR\_004343.  
/FtId=VAR\_004344.  
/FtId=VAR\_004345.  
/FtId=VAR\_004346.  
/FtId=VAR\_004347.  
/FtId=VAR\_004348.  
/FtId=VAR\_004349.  
/FtId=VAR\_004350.  
/FtId=VAR\_004351.  
/FtId=VAR\_004352.  
/FtId=VAR\_004353.  
/FtId=VAR\_004354.  
/FtId=VAR\_004355.  
/FtId=VAR\_004356.  
/FtId=VAR\_004357.  
/FtId=VAR\_004358.  
/FtId=VAR\_004359.  
/FtId=VAR\_004360.  
/FtId=VAR\_004361.  
/FtId=VAR\_004362.  
/FtId=VAR\_004363.  
/FtId=VAR\_004364.  
/FtId=VAR\_004365.  
/FtId=VAR\_004366.  
/FtId=VAR\_004367.  
/FtId=VAR\_004368.  
/FtId=VAR\_004369.  
/FtId=VAR\_004370.  
/FtId=VAR\_004371.  
/FtId=VAR\_004372.  
/FtId=VAR\_004373.  
/FtId=VAR\_004374.  
/FtId=VAR\_004375.  
/FtId=VAR\_004376.  
/FtId=VAR\_004377.  
/FtId=VAR\_004378.  
/FtId=VAR\_004379.  
/FtId=VAR\_004380.  
/FtId=VAR\_004381.  
/FtId=VAR\_004382.  
/FtId=VAR\_004383.  
/FtId=VAR\_004384.  
/FtId=VAR\_004385.  
/FtId=VAR\_004386.  
/FtId=VAR\_004387.  
/FtId=VAR\_004388.  
/FtId=VAR\_004389.  
/FtId=VAR\_004390.  
/FtId=VAR\_004391.  
/FtId=VAR\_004392.  
/FtId=VAR\_004393.  
/FtId=VAR\_004394.  
/FtId=VAR\_004395.  
/FtId=VAR\_004396.  
/FtId=VAR\_004397.  
/FtId=VAR\_004398.  
/FtId=VAR\_004399.  
/FtId=VAR\_004400.  
/FtId=VAR\_004401.  
/FtId=VAR\_004402.  
/FtId=VAR\_004403.  
/FtId=VAR\_004404.  
/FtId=VAR\_004405.  
/FtId=VAR\_004406.  
/FtId=VAR\_004407.  
/FtId=VAR\_004408.  
/FtId=VAR\_004409.  
/FtId=VAR\_004410.  
/FtId=VAR\_004411.  
/FtId=VAR\_004412.  
/FtId=VAR\_004413.  
/FtId=VAR\_004414.  
/FtId=VAR\_004415.  
/FtId=VAR\_004416.  
/FtId=VAR\_004417.  
/FtId=VAR\_004418.  
/FtId=VAR\_004419.  
/FtId=VAR\_004420.  
/FtId=VAR\_004421.  
/FtId=VAR\_004422.  
/FtId=VAR\_004423.  
/FtId=VAR\_004424.  
/FtId=VAR\_004425.  
/FtId=VAR\_004426.  
/FtId=VAR\_004427.  
/FtId=VAR\_004428.  
/FtId=VAR\_004429.  
/FtId=VAR\_004430.  
/FtId=VAR\_004431.  
/FtId=VAR\_004432.  
/FtId=VAR\_004433.  
/FtId=VAR\_004434.  
/FtId=VAR\_004435.  
/FtId=VAR\_004436.  
/FtId=VAR\_004437.  
/FtId=VAR\_004438.  
/FtId=VAR\_004439.  
/FtId=VAR\_004440.  
/FtId=VAR\_004441.  
/FtId=VAR\_004442.  
/FtId=VAR\_004443.  
/FtId=VAR\_004444.  
/FtId=VAR\_004445.  
/FtId=VAR\_004446.  
/FtId=VAR\_004447.  
/FtId=VAR\_004448.  
/FtId=VAR\_004449.  
/FtId=VAR\_004450.  
/FtId=VAR\_004451.  
/FtId=VAR\_004452.  
/FtId=VAR\_004453.  
/FtId=VAR\_004454.  
/FtId=VAR\_004455.  
/FtId=VAR\_004456.  
/FtId=VAR\_004457.  
/FtId=VAR\_004458.  
/FtId=VAR\_004459.  
/FtId=VAR\_004460.  
/FtId=VAR\_004461.  
/FtId=VAR\_004462.  
/FtId=VAR\_004463.  
/FtId=VAR\_004464.  
/FtId=VAR\_004465.  
/FtId=VAR\_004466.  
/FtId=VAR\_004467.  
/FtId=VAR\_004468.  
/FtId=VAR\_004469.  
/FtId=VAR\_004470.  
/FtId=VAR\_004471.  
/FtId=VAR\_004472.  
/FtId=VAR\_004473.  
/FtId=VAR\_004474.  
/FtId=VAR\_004475.  
/FtId=VAR\_004476.  
/FtId=VAR\_004477.  
/FtId=VAR\_004478.  
/FtId=VAR\_004479.  
/FtId=VAR\_004480.  
/FtId=VAR\_004481.  
/FtId=VAR\_004482.  
/FtId=VAR\_004483.  
/FtId=VAR\_004484.  
/FtId=VAR\_004485.  
/FtId=VAR\_004486.  
/FtId=VAR\_004487.  
/FtId=VAR\_004488.  
/FtId=VAR\_004489.  
/FtId=VAR\_004490.  
/FtId=VAR\_004491.  
/FtId=VAR\_004492.  
/FtId=VAR\_004493.  
/FtId=VAR\_004494.  
/FtId=VAR\_004495.  
/FtId=VAR\_004496.  
/

Best Local Similarity 90.3%; Pred. No. 2.6e-71;  
Matches 205; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

QY 2 DKTHTCPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
Db 64 DTPPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVQFKWYVD 123

QY 62 GVEVHNAKTPREQNSYRVSVLTVLHODWLNKCKYKCKVSNKALPAPIEKTISKAK 121  
Db 124 GVQVHNAKTPREQNSYRVSVLTVLHODWLNKCKYKCKVSNKALPAPIEKTISKYK 183

QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181  
Db 184 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 243

QY 182 DGSFFLYSKLTVDKSRWQQGNFSCVMHHEALHNHYTQKSLSLSPGK 228  
Db 244 DGSFFLYSKLTVDKSRWQQGNFSCVMHHEALHNHYTQKSLSLSPGK 290

RESULT 5  
GC\_RABIT STANDARD; PRT; 323 AA.  
AC P01870;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA CHAIN C REGION.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84030930; PubMed=6313520;  
RA Bernstein K.E., Alexander C.B., Mage R.G.;  
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant  
RT F-1 haplotype."  
RL Immunogenetics 18:387-397(1983).  
RN [2]  
RP SEQUENCE OF 1-128.  
RX MEDLINE=76135469; PubMed=1243651;  
RA Pratt D.M., Mole L.E.;  
RT "Sequence studies on the constant region of the Fd sections of rabbit  
RT immunoglobulin G of different allotype."  
RL Biochem. J. 151:337-349(1975).  
RN [3]  
RP SEQUENCE OF 88-266 FROM N.A.  
RX MEDLINE=83299917; PubMed=6193512;  
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;  
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma  
RT heavy chain and identification of two genomic C gamma genes."  
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).  
RN [4]  
RP SEQUENCE OF 132-161.  
RX MEDLINE=70110015; PubMed=5461106;  
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;  
RT "Sequence studies of the Fd section of the heavy chain of rabbit  
RT immunoglobulin G."  
RL Biochem. J. 116:249-259(1970).  
RN [5]  
RP SEQUENCE OF 129-131 AND 155-322.  
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;  
RL (In) Killander J. (eds.);  
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,  
RL Stockholm (1967).  
CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,  
CC 104-THR, AND REF.5 THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15  
CC MARKERS AND REF.5 THE E15 MARKER.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M16426; AAA31289.1; -;  
DR PIR; A02161; GHRB.  
DR INTERPRO; IPR000495; -;  
DR INTERPRO; IPR003006; -;  
DR PFAM; PF00047; ig; 3;  
DR PROSITE; PS00290; IG\_MHC; 1;  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1 1  
FT VARIANT 104 104 T -> M (IN D11 MARKER).  
FT VARIANT 185 185 T -> A (IN E15 MARKER).  
FT CONFLICT 48 48 N -> E (IN REF. 2).  
FT CONFLICT 71 71 V -> VPV (IN REF. 2).  
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).  
FT CONFLICT 173 173 N -> D (IN REF. 5).  
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).  
FT CONFLICT 201 201 N -> D (IN REF. 5).  
FT CONFLICT 218 218 Q -> E (IN REF. 5).  
FT CONFLICT 233 233 E -> Q (IN REF. 5).  
FT CONFLICT 246 246 N -> D (IN REF. 5).  
FT CONFLICT 256 256 E -> G (IN REF. 5).  
FT CONFLICT 260 260 N -> D (IN REF. 5).  
FT CONFLICT 266 266 N -> D (IN REF. 5).  
FT CONFLICT 280 280 Y -> W (IN REF. 5).  
FT CONFLICT 284 284 N -> S (IN REF. 5).  
SQ SEQUENCE 323 AA; 35404 MW; 698AA118D579A8B CRC64;

Query Match 40.9%; Score 918.5; DB 1; Length 323;  
Best Local Similarity 71.7%; Pred. No. 6.7e-57;  
Matches 167; Conservative 29; Mismatches 32; Indels 5; Gaps 2;

QY 1 MDKT---HTC--PPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 55  
Db 91 VDKTVPSTCTSKPTCPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 150

QY 56 FNNYDGVVHNAKTPREQNSYRVSVLTVLHODWLNKCKYKCKVSNKALPAPIEK 115  
Db 151 FTWYINNEQVTRAPPLREQQFNSTIRVSTLPTITHODWLNKCKYKCKVSNKALPAPIEK 210

QY 116 TISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 175  
Db 211 TISKARGQPLEPKVYTMGPPREELSSRSVSLTCLMNGYFSDISVEWEKNGKAEDNKTT 270

QY 176 PPVLDSGDSFFLYSKLTVDKSRWQQGNFSCVMHHEALHNHYTQKSLSLSPGK 228  
Db 271 PAVLDSGDSFFLYSKLTVDKSRWQQGNFSCVMHHEALHNHYTQKSLSRSPGK 323

RESULT 6  
GC2\_CAVPO STANDARD; PRT; 329 AA.  
ID GC2\_CAVPO  
AC P01862;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2 CHAIN C REGION.  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
RN [1]  
RP SEQUENCE OF 1-3.  
RA Trischmann T.M.;  
RL Submitted (APR-1975) to the PIR data bank.  
RN [2]  
RP SEQUENCE OF 4-68.  
RX MEDLINE=71058471; PubMed=5538606;  
RA Birshstein B.K., Hussain Q.Z., Cebra J.J.;  
RT "Structure of heavy chain from strain 13 guinea pig  
RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the

RT half-cystine joining heavy and light chains.";  
 RL Biochemistry 10:18-25(1971).  
 RN [3]  
 RP SEQUENCE OF 69-133 AND 312-329.  
 RX MEDLINE=71058486; PubMed=5538616;  
 RA Turner K.J., Cebra J.J.;  
 RT "Structure of heavy chain from strain 13 guinea pig  
 RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal  
 RT and hinge region cyanogen bromide fragments.";  
 RL Biochemistry 10:9-17(1971).  
 RN [4]  
 RP SEQUENCE OF 134-226.  
 RX MEDLINE=75036072; PubMed=4429665;  
 RA Tracey D.E., Cebra J.J.;  
 RT "Primary structure of the CH2 homology region from guinea pig IgG2  
 RT antibodies";  
 RL Biochemistry 13:4796-4803(1974).  
 RN [5]  
 RP SEQUENCE OF 227-311.  
 RX MEDLINE=75036073; PubMed=4609467;  
 RA Trischmann T.M., Cebra J.J.;  
 RT "Primary structure of the CH3 homology region from guinea pig IgG2  
 RT antibodies";  
 RL Biochemistry 13:4804-4811(1974).  
 RN [6]  
 RP DISULFIDE BONDS.  
 RX MEDLINE=71058474; PubMed=4922544;  
 RA Oliveira B., Lamm M.E.;  
 RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";  
 RL Biochemistry 10:26-31(1971).  
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN  
 CC 13 INBRED GUINEA PIGS.  
 DR PIR: A02151; G2GP.  
 DR INTERPRO: IPR000495;  
 DR INTERPRO: IPR003006;  
 DR PFAM: PF00047; Ig; 3.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
 FT NON\_TER 1 1  
 FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 28 79  
 FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 142 202  
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).  
 FT DISULFID 248 308  
 SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 39.6%; Score 889; DB 1; Length 329;  
 Best Local Similarity 72.3%; Pred. No. 7.5e-55;  
 Matches 162; Conservative 24; Mismatches 36; Indels 2; Gaps 1;  
 Qy 6 TCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGEV 65  
 Db 106 TCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGEV 165  
 Qy 66 HNAKTPREEQYNSTYRVSVLTIVLHODWLNKGEYKCKVSKNKPAPLPIETISKAKGQPR 125  
 Db 166 GNAETKPRVEQYNTFRVESVLPVHOQDWMKGEYKCKVSKNKPAPLPIETISKAKGQPR 225  
 Qy 126 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYTKTTPPVLDSGDG 183  
 Db 226 MPDYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYTKTTPPVLDSGDG 285  
 Qy 184 SFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 227  
 Db 286 SYFLYSLKLTVDKSAWDQGTQVTCSCVMHEALHNHYTQKSLSLSPG 329

RESULT 7  
 GC3\_MOUSE

GC3\_MOUSE STANDARD; PRT; 329 AA.  
 AC P22436;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-3 CHAIN C REGION, SECRETED FORM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85027161; PubMed=6092053;  
 RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,  
 RA Tucker P.W., Blattner F.R.;  
 RT "Structural analysis of the murine IgG3 constant region gene.";  
 RL EMBO J. 3:2041-2046(1984).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 DR EMBL: J00451; -: NOT\_ANNOTATED\_CDS.  
 DR PIR: B02156; G3MSC.  
 DR INTERPRO: IPR000495;  
 DR INTERPRO: IPR003006;  
 DR PFAM: PF00047; Ig; 3.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW Immunoglobulin C region; Glycoprotein; Transmembrane;  
 FT NON\_TER 1 97  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 113 HINGE.  
 FT DOMAIN 114 223 CH2.  
 FT DOMAIN 224 327 CH3.  
 SQ SEQUENCE 329 AA; 36228 MW; F4582717A182BAD6 CRC64;

Query Match 37.6%; Score 845.5; DB 1; Length 329;  
 Best Local Similarity 67.0%; Pred. No. 7.7e-52;  
 Matches 150; Conservative 34; Mismatches 37; Indels 3; Gaps 1;  
 Qy 8 PP---CPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGE 64  
 Db 106 PPGSCPPGNTLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGE 165  
 Qy 65 VHNKTKPREQYNSTYRVSVLTIVLHODWLNKGEYKCKVSKNKPAPLPIETISKAKGQPR 124  
 Db 166 VHTATQPREAQYNSTFRVVSALPIQHDWMKGEYKCKVSKNKPAPLPIETISKAKGQPR 225  
 Qy 125 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYTKTTPPVLDSGDG 184  
 Db 226 QTPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYTKTTPPVLDSGDG 285  
 Qy 185 FFLLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 228  
 Db 286 YFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPG 329

RESULT 8  
 GC3\_MOUSE STANDARD; PRT; 333 AA.  
 AC P20761;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE IG GAMMA-2B CHAIN C REGION.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.



```

[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR INTERPRO: PS0018; PS0018.
DR INTERPRO: IPR000495; -.
DR PFAM: PF00047; Ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match
Best Local Similarity 37.3%; Score 838; DB 1; Length 333;
Matches 148; Conservative 33; Mismatches 42; Indels 0; Gaps 0;

QY 6 TCPCPAPELLGGPSVFLPPKPKDLMISRTPEVTCVVDVSHEDPEVKFNYYDGV 65
DB 111 TCHKCPPELLGGPSVFIAPPKPKDILLISQNAKVTCTVVDVSEEPDVQFSWFVNVEV 170
QY 66 HNAKTKPREQYNSTRYVVSALPIQHDWNGREYKCKVSNKALPAPIETKISKAKG 125
DB 171 HTAQTPREQYNSTRYVVSALPIQHDWNGREYKCKVSNKALPAPIETKISKAKG 230
QY 126 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGSF 185
DB 231 KPQVYVNGPTEQLTEQTSVLTCLTSGFLPNDIGVEMTSNGHKYKNTPEVWDSDGSF 290
QY 186 FLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 228
DB 291 FMYSKLVNRSWRDRAFPVCSVHVEGLHHHVEKSISSPPGK 333

RESULT 9
GC3M_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-3 CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Weis J.A.; Word C.J.; Rimm D.; Der-Balan G.P.; Martinez H.M.;
RA Tucker P.W.; Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
[2]
RN SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M.; Clayton L.; Rogers J.; Robertson S.; Kettman J.;
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: J00451; AAB59655.1; -.
DR EMBL: V01526; CAA24767.1; ALT_SEQ.
DR PIR: A02155; G3MSM.
DR INTERPRO: IPR000495; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; Ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DOMAIN 98 113 CH1.
FT DOMAIN 114 223 HINGE.
FT DOMAIN 224 327 CH2.
FT DOMAIN 327 362 CH3.
FT TRANSMEM 345 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 333 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match
Best Local Similarity 37.2%; Score 835.5; DB 1; Length 398;
Matches 149; Conservative 34; Mismatches 39; Indels 3; Gaps 1;

QY 8 PP---CPAPPELLGGPSVFLPPKPKDLMISRTPEVTCVVDVSHEDPEVKFNYYDGV 64
DB 106 PGSSCPPGILGGPSVFIAPPKPKDLMISLTPKVTCTVVDVSEDDPDVHVSFVDNKE 165
QY 65 VHAKTTPREQYNSTRYVVSALPIQHDWNGREYKCKVSNKALPAPIETKISKAKG 124
DB 166 VHTAWTPREQYNSTRYVVSALPIQHDWNGREYKCKVSNKALPAPIETKISKAKG 225
QY 125 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGS 184
DB 226 QTPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLDSDGT 285
QY 185 FFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGKE 229
DB 286 YFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGKE 330

RESULT 10
GCL_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR: PS0017; PS0017.
DR INTERPRO: IPR000495; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; Ig; 3.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
```

```
FT NON_TER 1 1 CH1.
FT DOMAIN 1 97
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 36.4%; Score 818.5; DB 1; Length 326;
Best Local Similarity 63.1%; Pred. No. 5.6e-50;
Matches 142; Conservative 41; Mismatches 35; Indels 7; Gaps 2;

QY 7 CPPCPAPPELLGG---PSVLEPPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 63
Db 106 CKPC-----ICGSEVSSVFIFPPKPKDVLTLTPKVCVVVDISQDDPEVHFVFWFDDV 161

QY 64 EVHNAKTPREEQNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAKGQ 123
Db 162 EVHTAQTPPEEQNSTFRSVSELPILHODWLNKTRFCKVTSAAFPSPIEKTIKSGEGR 221

QY 124 PREQVYTLPSRELTKNOVSLCLVKGYFSPDAVWESNGQPENNYKTPPVLDSDG 183
Db 222 TVPHVYMTSKTEMTQNEVSIICMVKGYFPDIIYEWQMGQPQENYKNTPTPTMDTG 281

QY 184 SFYLSKLTVDKSRWQGNVFCSSVMHEALHNHYTKSLSPGK 228
Db 282 SYFLSKLVNKKERKQGGNTFTCSVLHGLNHNHTEKSLSPGK 326

RESULT 11
GCL_MOUSE STANDARD; PRT; 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-1 CHAIN C REGION.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RN SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RX MEDLINE=80202559; PubMed=6769752;
RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
cloned in a bacterial plasmid.";
RL Gene 9:87-97(1980).
RN [3]
RN SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=80012837; PubMed=113776;
RA Rogers J., Clarke P., Salsner W.;
RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
heavy chain.";
RL Nucleic Acids Res. 6:3305-3321(1979).
RN [4]
RN SEQUENCE (MYELOMA PROTEIN MOPC 21).
```

```
RX MEDLINE=78242288; PubMed=98524;
RA Adetugbo K.;
RT "Evolution of immunoglobulin subclasses. Primary structure of a
murine myeloma gammal chain.";
RL J. Biol. Chem. 253:6068-6075(1978).
RN [5]
RN DISULFIDE BONDS (MOPC 21).
RX MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RL Biochem. J. 126:837-850(1972).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR EMBL; V00793; CAA24175.1; -
DR EMBL; V00795; CAA24176.1; -
DR PIR; A02159; GIMS.
DR MGD; MGI:96446; Igh-4.
DR INTERPRO; IPR000495; -
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; Iy; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
FT NON_TER 1 97
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 36.4%; Score 817.5; DB 1; Length 324;
Best Local Similarity 62.5%; Pred. No. 6.6e-50;
Matches 140; Conservative 44; Mismatches 35; Indels 5; Gaps 2;

QY 7 CPP--CPAPELGGSVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 64
Db 104 CKPICVPEV---SSVFIFPPKPKDVLTLTPKVCVVVDISKDDPEVQFSWFVDDVE 160

QY 65 VHNAKTPREEQNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAGQP 24
Db 161 VHTAQTPREEQNSTFRSVSELPIMHODWLNKKEFCRVNSAFAPIETIKTKGRP 220

QY 125 REPOVYTLPPSRDELTKNOVSLTCLVKGYFSPDAVWESNGQPENNYKTPPVLDSDGS 184
Db 221 KAPQVYTIPTPPKEQMAKDKVSLTCDITDFFPEDITVEMQWNGQPAENYKNTPTMTNGS 280

QY 185 FFYLSKLTVDKSRWQGNVFCSSVMHEALHNHYTKSLSPGK 228
Db 281 YFVYLSKLVNKKERKQGGNTFTCSVLHGLNHNHTEKSLSPGK 324
```

RESULT 12  
GC1M\_MOUSE STANDARD; PRT; 393 AA.  
AC P01869;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
RN [1]  
RN [1]  
SEQUENCE FROM N.A.  
RX MEDLINE=80045036; PubMed=115593;  
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
RA Takahashi N., Mano Y.;  
RT Cloning and complete nucleotide sequence of mouse immunoglobulin  
RT gamma 1 chain gene.;  
RL Cell 18:559-568(1979).  
RN [2]  
RN [2]  
SEQUENCE OF 323-393 FROM N.A.  
RX MEDLINE=82197626; PubMed=6804950;  
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;  
RT "mRNA for surface immunoglobulin gamma chains encodes a highly  
RT conserved transmembrane sequence and a 28-residue intracellular  
RT domain";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).  
RN [3]  
RN [3]  
SEQUENCE OF 323-366 FROM N.A.  
RX MEDLINE=82115295; PubMed=6799207;  
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,  
RA Eisenberg D., Wall R.;  
RT "Gene segments encoding transmembrane carboxyl termini of  
RT immunoglobulin gamma chains";  
RL Cell 26:19-27(1981).  
RN [4]  
RN [4]  
SEQUENCE OF 1-44 FROM N.A.  
RX MEDLINE=8222190; PubMed=6283537;  
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;  
RT "Nucleotide sequences of gene segments encoding membrane domains of  
RT immunoglobulin gamma chains";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).  
CC -1- ALTERNATIVE PRODUCTS: CHAIN LINES PRODUCING IGG CONTAIN TWO MRNA  
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED  
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-  
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED  
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND  
CC SEGMENT OF MU CHAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; V00793; CAA241172.1; -;  
CC EMBL; V00793; CAA241173.1; -;  
CC EMBL; V00793; CAA241174.1; -;  
CC PIR; B02159; GIMSM.  
CC MGD; MGI:96446; IGH-4.  
CC INTERPRO: IPR000495; -;  
CC INTERPRO: IPR003006; -;  
CC PFAM; PF00047; ig; 3.  
CC PROSITE; PS00290; IG\_MHC; 1.  
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
CC Alternative splicing; Transmembrane.  
CC  
CC NON\_TER 1 1  
CC DOMAIN 1 97 CH1.  
CC DOMAIN 98 110 HINGE.  
CC DOMAIN 111 217 CH2.  
CC DOMAIN 218 324 CH3.

FT DISULFID 27 82  
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 138 138 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 174 174  
FT DISULFID 244 302  
FT TRANSMEM 340 357  
FT DOMAIN 358 393 POTENTIAL.  
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7AICE27 CRC64;  
CYTOPLASMIC (POTENTIAL).  
Query Match 36.2%; Score 812.5; DB 1; Length 393;  
Best Local Similarity 62.3%; Pred. No. 1.8e-49;  
Matches 139; Conservative 44; Mismatches 35; Indels 5; Gaps 2;  
QY 7 CPP--CPAPELLGGPSVFLFPPPKDPTLMTISRTPEVTCVVYVDSHEDPEYKFNMYVDGVE 64  
Db 104 CKPKICTVPEV---SSVFIPPKPKDVLITLTPKVTCTVVVDISKDDPEVQFSWFVDVDE 160  
QY 65 VHNAKTKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSKNKPALPAPIETISKAKQP 124  
Db 161 VHTAQTQPREEQFNSTFRSVSELPIMHODWLNKGYCKVSKNKPALPAPIETISKAKGRP 220  
QY 125 REPQVYTLPPSRDELTKNOVSLCLVKGFYPSDIAVWESNGOPENNYKTPPVLDSDGS 184  
Db 221 KAFQVYTIPTPPKQMAKDKVSLCTMTDFPEDITVEWQNGQPAENYKNTQPIINTNGS 280  
QY 185 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPG 227  
Db 281 YFVYSKLVNQSKEAGNTEFCSVLHGLHNHHTKSLSHSPG 323  
RESULT 13  
GCC\_RAT STANDARD; PRT; 329 AA.  
AC P20762;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG GAMMA-2C CHAIN C REGION.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
RN [1]  
RN [1]  
SEQUENCE FROM N.A.  
RX MEDLINE=88166903; PubMed=3127222;  
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;  
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant  
RT region cDNA: extensive homology to mouse gamma 3";  
RL Eur. J. Immunol. 18:317-319(1988).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X07189; CAA30169.1; -;  
CC PIR; S00847; S00847.  
CC INTERPRO: IPR000495; -;  
CC INTERPRO: IPR003006; -;  
CC PFAM; PF00047; ig; 3.  
CC PROSITE; PS00290; IG\_MHC; 1.  
CC Immunoglobulin domain; Immunoglobulin C region.  
CC  
CC NON\_TER 1 1  
CC DOMAIN 1 97 CH1.  
CC DOMAIN 98 113 HINGE.  
CC DOMAIN 114 222 CH2.  
CC DOMAIN 223 329 CH3.

```
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 36.0%; Score 809.5; DB 1; Length 329;
Best Local Similarity 63.8%; Pred. No. 2.4e-49;
Matches 143; Conservative 41; Mismatches 37; Indels 3; Gaps 1;

Qy 8 PP---CPAPELGGPSVFLFPKPKDTLMISRTPTVTCVVVDVSHEDPEVKFNWVDGVE 64
Db 106 PPTDCSCDDNLGRPSVFIFPPKPKDILMTLPKVTCTVVVDVSEEDPVQFSEFVDNR 165

Qy 65 VHNATKPREQYNSTYRVSVLTVLVHODWLNKGYCKVSNKALPAPIETISKAKGQP 124
Db 166 VFTAQTPHEQLNGTFRVSVTLTHQHDWMSGKFEKCKVANKDLPSPKPRGKA 225

Qy 125 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEGSPENNYKTPPVLDSDGS 184
Db 226 RTPQYITIPPREQMSKKNVSLTCMTVFYFASISVERNGELEDQYKNTLPVLDSDS 285

Qy 185 FFLYSKLTVDKSRMOOGNVFSCVMHEALHNHYTKSLSPGK 228
Db 286 YFLYSKLSVDTDWNRGDIYTCVSVVHEALHNHYTKSLSPGK 329

RESULT 14
GCAA_MOUSE STANDARD; PRT; 330 AA.
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, A ALLELE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE-8107654; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RT Balb/c gamma 2a heavy chain messenger RNA.";
RL Nucleic Acids Res. 8:3143-3155(1980).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE-81198976; PubMed=6262729;
RA Yanawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
RT and evolution of heavy chain genes: further evidence for intervening
RT sequence-mediated domain transfer.";
RL Nucleic Acids Res. 9:1365-1381(1981).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE-81223894; PubMed=6787604;
RA Ollo R., Auffray C., Mochamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
RT suggests that exons can be exchanged between genes in a multigenic
RT family.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
[4]
RN MYELOMA PROTEIN MOPC 173.
RX MEDLINE-74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
RT immunoglobulin:amino-acid sequence of the FC fragment. Implications
RT for the evolution of immunoglobulin structure and function.";
RL Eur. J. Biochem. 43:423-435(1974).
[5]
```

```
RP DISULFIDE BONDS.
RX MEDLINE-73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
RT immunoglobulin. Identification of the disulfide bridges.";
RL Eur. J. Biochem. 30:452-462(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00798; CAA24178.1; -.
DR PIR; A02152; G2MSA.
DR INTERPRO; IPR000495; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; Ig; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT MOD_RES 330 330
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 36.0%; Score 809; DB 1; Length 330;
Best Local Similarity 65.2%; Pred. No. 2.6e-49;
Matches 146; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

Qy 7 CPP---CPAPELGGPSVFLFPKPKDTLMISRTPTVTCVVVDVSHEDPEVKFNWVDGVE 64
Db 107 CPPCKPAPNLGGPSVFIFPPKINKDVLMLSLSPITCVVVDSEDDPDVQLSWFNWVE 166

Qy 65 VHNATKPREQYNSTYRVSVLTVLVHODWLNKGYCKVSNKALPAPIETISKAKGQP 124
Db 167 VHTAQTPHEQYNSTLRVVSALPIQHDWMSGKFEKCKVANKDLPAPIETISKPKGSV 226

Qy 125 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEGSPENNYKTPPVLDSDGS 184
Db 227 RAPQVYVLPPEEEMTKKQVTLTCMTVFDPEDYVETNNGKTELKNTPEVLDSDGS 286

Qy 185 FFLYSKLTVDKSRMOOGNVFSCVMHEALHNHYTKSLSPGK 228
Db 287 YFLYSKLVKKNVVERNSYSCSVVHEGLHNHTTKSFRTPGK 330

RESULT 15
GCAA_MOUSE STANDARD; PRT; 399 AA.
AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION, MEMBRANE-BOUND FORM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE-82222190; PubMed=6283537;
RA Yanawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
```

CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA  
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED  
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-  
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END. ENCODED  
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND  
CC SEGMENT OF MU CHAINS.  
CC  
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE  
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF  
CC THE A ALLELE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; J00471; AAB59661.1; ALT\_INIT.  
DR PIR; A02154; G2MSAM.  
DR MGI; 96443; IGH-1.  
DR INTERPRO: IPR000495; -  
DR INTERPRO: IPR003006; -  
DR PFAM; PF00047; 19; 3.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
KW Transmembrane; Alternative splicing.  
FT NON\_TER 1 1  
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 27 82  
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 144 204  
FT DISULFID 250 308  
FT TRANSMEM 346 363  
FT DOMAIN 364 399 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;

Query Match 35.8%; Score 804; DB 1; Length 399;  
Best Local Similarity 65.0%; Pred. NO. 7.2e-49;  
Matches 145; Conservative 30; Mismatches 46; Indels 2; Gaps 1;

QY 7 CPP--CPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 64  
Db 107 CPPCKCPAPNLLGGPSVFIKPKIKDVLMSLSPIVTCVVVDVSEDDPDYQISFEVNNVE 166  
QY 65 VHNAKTKPREEQNSTYRVYSVLTVLHODWLNKGKCKVKSNKALPAPIEKTISKAKGQP 124  
Db 167 VHTAQQTREDYNSLRVVSALPIQHDWMSGKFKACKNNKDLPAPIERTISKPGSV 226  
QY 125 REPQVYTLPPSRDLTKNQVSLTCLVKGYFSPDAVWESNGQPENNYKTTTPVLDSDGS 184  
Db 227 RAPQVYVLPPEEEMTKQVTLTCMTDFWPDYIYVETWNTNGKTELNYKNTPEVLDSDGS 286  
QY 185 FFYLSKLTVDKSRQOQNVFSCSMHEALHNHYTKSLSPG 227  
Db 287 YFMYSKLRVEKKNVERNYSYCSVHVEGLHNHHTTKFSRTPG 329

Search completed: March 1, 2001, 09:17:49  
Job time: 267 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:15:49 ; Search time 79.26 Seconds  
(without alignments)  
343.530 Million cell updates/sec

Title: US-09-389-782A-8  
Perfect score: 2246  
Sequence: 1 MDKTHTCPCPAPELLGGPS.....QKNATHDNCISGNSESPQK 401

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 6790655 residues  
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_66:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	54.9	255	4 S31866	Ig gamma-1 chain C
2	1233	54.9	330	1 GHUU	Ig gamma-1 chain C
3	1227	54.6	374	2 S69339	Ig heavy chain V r
4	1227	54.6	374	2 S72664	Ig heavy chain V r
5	1180	52.5	234	2 PT0207	Ig gamma chain C r
6	1146	51.0	377	2 A23511	Ig gamma-3 chain C
7	1144	50.9	377	2 A60764	Ig gamma-3 chain C
8	1142.5	50.9	326	1 G2HU	Ig gamma-2 chain C
9	1135	50.5	327	1 G4HU	Ig gamma-4 chain C
10	1121	49.9	289	1 G3HUI	Ig gamma-3 heavy c
11	918.5	40.9	323	1 GHRB	Ig gamma chain C r
12	906.5	40.4	328	2 I47160	Ig gamma 2b chain
13	906.5	40.4	328	2 I47159	Ig gamma 2a chain
14	903.5	40.2	277	2 I47162	Ig gamma 4 chain C
15	889	39.6	329	1 G2GP	Ig gamma-2 chain C
16	885.5	39.4	328	2 I47158	Ig gamma 1 chain c
17	878.5	39.1	328	2 I47161	Ig gamma 3 chain c
18	855.5	38.1	470	2 S22080	Ig heavy chain pre
19	846	37.7	308	2 S30554	Ig heavy chain C r
20	846	37.7	472	2 S31459	Ig gamma-1 chain -
21	845.5	37.6	329	1 G3MSC	Ig gamma-3 chain C
22	838	37.3	333	2 PS0018	Ig gamma-2b chain
23	835.5	37.2	398	1 G3MSM	Ig gamma-3 chain C
24	827.5	36.8	444	2 PC4436	monoclonal antibod
25	818.5	36.4	326	2 PS0017	Ig gamma-1 chain C
26	817.5	36.4	324	1 G1MS	Ig gamma-1 chain C
27	812.5	36.2	393	1 G1MSM	Ig gamma-1 chain C
28	809.5	36.0	329	2 S00847	Ig gamma-2c chain
29	809	36.0	330	1 G2MSA	Ig gamma-2a chain

30	809	36.0	469	2 S37483	Ig gamma-2a chain
31	804	35.8	399	1 G2MSAM	Ig gamma-2a chain
32	802	35.7	335	1 G2MSAB	Ig gamma-2a chain
33	794	35.4	446	2 S40295	Ig gamma-2a chain
34	785.5	35.0	322	2 PS0019	Ig gamma-2b chain
35	779	34.7	336	1 G2MS11	Ig gamma-2b chain
36	779	34.7	474	2 S25057	Ig gamma-2b chain
37	774	34.5	405	1 G2MSBM	Ig gamma-2b chain
38	764	34.0	327	2 S06611	Ig gamma-2 chain C
39	757	33.7	475	2 S01321	Ig gamma-2b chain
40	707	31.5	180	2 I46732	Ig gamma heavy cha
41	577.5	25.7	249	2 S69340	Ig heavy chain VHI
42	574.5	25.6	218	2 A36040	Ig heavy chain V-I
43	571	25.4	152	2 S14236	Ig gamma-1 chain C
44	395.5	17.6	572	2 B46529	Ig Y heavy chain (
45	362	16.1	343	2 S25644	Ig mu chain C regi

## ALIGNMENTS

### RESULT 1

S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli  
C:Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000  
C:Accession: S31866

R:Filpula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene produc

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FIL>

A:Cross-references: EMBL:X70421; MID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: Immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor ,

F:23-255/Region: human Ig gamma-1 chain C region

Query Match 54.9%; Score 1233; DB 4; Length 255;  
Best Local Similarity 100.0%; Pred. No. 9.6e-72;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKHTCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVYDVSHEDPEVFNWVD 61

Db 29 DKHTCPCPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVYDVSHEDPEVFNWVD 88

QY 62 GVEVHNKATKPREQYNTVYRVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAK 121

Db 89 GVEVHNKATKPREQYNTVYRVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAK 148

QY 122 GQPREQVYITLPPSRDELTKNQVSLTCLIVKGFPYSDIAVWESNGQPENNYKTPPPVLD 181

Db 149 GQPREQVYITLPPSRDELTKNQVSLTCLIVKGFPYSDIAVWESNGQPENNYKTPPPVLD 208

QY 182 DGSFPLYSKLTVDKSRQOQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228

Db 209 DGSFPLYSKLTVDKSRQOQGNVFCSCVMHEALHNHYTQKSLSLSPGK 255

### RESULT 2

GHUU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene

A:Reference number: A93433; MUID:82274238

A:Accession: A93433

A:Molecule type: DNA  
A:Residues: 1-330 <ELL>  
A:Cross-references: EMBL:Z117370  
A:Note: this sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers,  
A:Note: Lys-330 is removed after translation  
R:Harris, L.J.  
submitted to the EMBL Data Library, October 1992  
A:Reference number: S33904  
A:Accession: S36861  
A:Molecule type: DNA  
A:Residues: 2-330 <HAR>  
A:Cross-references: EMBL:Z117370  
R:Takahashi, N.; Oeda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.  
Cell 29, 671-679, 1982  
A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of a  
A:Reference number: S33887; MUID:83001943  
A:Accession: S33887  
A:Molecule type: DNA  
A:Residues: 88-113:235-330 <TAK>  
A:Cross-references: EMBL:Z117370  
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,  
Biochemistry 9, 3161-3170, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen  
A:Reference number: A90563; MUID:71064024  
A:Contents: myeloma protein Eu  
A:Accession: B90563  
A:Molecule type: protein  
A:Residues: 1-96,'R',98-135 <CUN>  
A:Note: this sequence has the G1m(3) marker, 97-Arg  
R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.  
Biochemistry 9, 3171-3181, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequen  
A:Reference number: A90564; MUID:71064025  
A:Contents: Eu  
A:Accession: A90564  
A:Molecule type: protein  
A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',240,  
A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met  
R:Ponstingl, H.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976  
A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie).  
igen Primärstruktur.  
A:Reference number: A91668; MUID:77070269  
A:Contents: myeloma protein Nie  
A:Accession: B91668  
A:Molecule type: protein  
A:Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27  
A:Note: this sequence has the G1m(17) and G1m(1) markers  
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983  
A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL  
A:Reference number: A91723; MUID:83289131  
A:Contents: myeloma protein KOL; disulfide bonds  
A:Accession: A91723  
A:Molecule type: protein  
A:Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH  
A:Note: this sequence has the G1m(3) and G1m(non-1) markers  
R:Gall, W.E.; Edelman, G.M.  
Biochemistry 9, 3188-3196, 1970  
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid  
A:Reference number: A90565; MUID:71064027  
A:Contents: annotation; disulfide bonds  
R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.  
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976  
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob  
enbromide cleavage products, and the disulfide bridges.  
A:Reference number: A91667; MUID:77070267  
A:Contents: annotation; disulfide bonds  
C:Genetics:  
A:Gene: GDB:IGH1  
A:Cross-references: GDB:120085; OMTM:147100  
A:Map position: 14q32.33-14q32.33  
A:Introns: 99/1; 114/1; 224/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (t  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:20-85/Domain: immunoglobulin homology <IM1>  
F:137-206/Domain: immunoglobulin homology <IM2>  
F:243-310/Domain: immunoglobulin homology <IM3>  
F:27-83,144-204,250-308/Disulfide bonds: #status experimental  
F:103/Disulfide bonds: interchain (to light chain) #status experimental  
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental  
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 54.9%; Score 1233; DB 1; Length 330;  
Best Local Similarity 100.08; Pred. No. 1.3e-71;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
DB 104 DKTHCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163  
QY 62 GVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 121  
DB 164 GVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 223  
QY 122 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGPSPDIAVEWESNGQPENNYKTTTPPVLD 181  
DB 224 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGPSPDIAVEWESNGQPENNYKTTTPPVLD 283

QY 182 DGSFELYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPCK 228  
DB 284 DGSFELYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPCK 330

## RESULT 3

Ig heavy chain V region precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 17-Mar-2000  
C:Accession: S69339  
R:Khamlich, A.A.; Aucouturier, P.; Preud'homme, J.L.; Coyne, M.  
Eur. J. Biochem. 229, 54-60, 1995  
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.  
A:Reference number: S69339; MUID:95262687  
A:Accession: S69339  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <KHA>  
A:Cross-references: EMBL:X81695  
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 54.6%; Score 1227; DB 2; Length 374;  
Best Local Similarity 99.18; Pred. No. 3.5e-71;  
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
DB 148 DKTHCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 207  
QY 62 GVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 121  
DB 208 GVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEKTISKAK 267

QY 122 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGPSPDIAVEWESNGQPENNYKTTTPPVLD 181  
DB 268 GQPREPQVYTLPPSRDEMTKNOVSLTCLVKGPSPDIAVEWESNGQPENNYKTTTPPVLD 327

QY 182 DGSFELYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPCK 228  
DB 328 DGSFELYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPCK 374



```

RESULT      4
S72664
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
C:Accession: S72664
R:hamichi, A.A.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72664
A:Accession: S72664
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match          54.6%; Score 1227; DB 2; Length 374;
Best Local Similarity 99.1%; Pred.No. 3.5e-71;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps

QY    2 DKTHTCPPCAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    148 DKTHTCPPCAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 207
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    62 GVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLGNGKEYCKVSNKALPAPIETISKAK 121
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    208 GVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLGNGKEYCKVSNKALPAPIETISKAK 267
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    122 GQREPOVYTLPPSRDELTKNQVSLTCLVKGYGSPIDIAVESNGQPENNYKTTPPVLD 181
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    268 GQREPOVYTLPPSRDEMTKNQVSLTCLVKGYGSPIDIAVESNGQPENNYKTTPPVLD 327
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    182 DGSFFLYSKITVDKSRWQQGNVFSCSVMHEALHNYHTOKSLSLSPGK 228
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    328 DGSFFLYSKITVDKSRWQQGNVFSCSVMHEALHNYHTOKSLSLSPGK 374
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT      5
PT0207
Ig gamma chain C region - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
C:Accession: PT0207
R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
Mol. Immunol. 28, 319-322, 1991
A>Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A:Reference number: PT0207; MUID:91287716
A:Accession: PT0207
A:Molecule type: mRNA
A:Residues: 1-234 <BHR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:48-117/Domain: immunoglobulin homology <IMM>

Query Match          52.5%; Score 1180; DB 2; Length 234;
Best Local Similarity 98.6%; Pred. No. 2.1e-68;
Matches 217; Conservative 1; Mismatches 2; Indels 0; Gaps

QY    2 DKTHTCPPCAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    15 DTTHTCPCAAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 74
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    62 GVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLGNGKEYCKVSNKALPAPIETISKAK 121
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    75 GVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLGNGKEYCKVSNKALPAPIETISKAK 134
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    122 GQREPOVYTLPPSRDELTKNQVSLTCLVKGYGSPIDIAVESNGQPENNYKTTPPVLD 181
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    135 GQREPOVYTLPPSRDELTKNQVSLTCLVKGYGSPIDIAVESNGQPENNYKTTPPVLD 194
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY    182 DGSFFLYSKITVDKSRWQQGNVFSCSVMHEALHNYHTOKS 221

```

```

Db      195 DGSFFLYSKLTVDKSRWQOGNVFCSCVMHEALHNHYTKS   234
|||||
RESULT    6
A23511
Ig gamma-3 chain C region (allotype G3m(b)) - human
C:Species: Homo sapiens (man)
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
C:Accession: A23511
R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.
Nucleic Acids Res. 14, 1779-1789, 1986
A>Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene:
A:Reference number: A23511; MUID:86148507
A:Accession: A23511
A:Molecule type: DNA
A:Residues: 1-377 <HUC>
A:Cross-references: GB:X03504; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056
C:Genetics:
A:Gene: IGHG3
A:Cross-references: GDB:l19339; OMIM:147120
A:Map position: 14q32.33-14q32.33
A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match          51.0%; Score 1146; DB 2; Length 377;
Best Local Similarity 92.5%; Pred. No. 5.1e-66;
Matches 210; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Qy  2 DKHTCCPCAPPELLGGPSVLFPPPKDPTLMISRTPEVTCTVVVDVSHEDPVEVKFNMYVD 61
|||
Db  151 DTPPPCPAPPELLGGPSVLFPPPKDPTLMISRTPEVTCTVVVDVSHEDPVEVKFNMYVD 210
|||

Qy  62 GVEYHNAKTTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCCKVSNKALPAPIEKTISKAK 121
|||
Db  211 GVEYHNAKTTPREQYNSTFRVSVLTVLHQDWLNGKEYKCCKVSNKALPAPIEKTISKTK 270
|||

Qy  122 QGPREPQVYTLPSPDELTKNQVSLTCLVGKFYPDSIAVEVESGNQPENNYKTTPPVLDLS 181
|||
Db  271 QGPREPQVYTLPSPGREEMTKNQVSLTCLVGKFYPDSIAVEVESGQPENNYNTTPPMLDS 330
|||

Qy  182 DGSFFLYSKLTVDKSRWQOGNVFCSCVMHEALHNHYTKQSLSLSPGK 228
|||||
Db  331 DGSFFLYSKLTVDKSRWQOGNIFCSVMHEALHNRFQTOKSLSLSPGK 377
|||||

RESULT    7
A60764
Ig gamma-3 chain C region, form LAT - human
C:Species: Homo sapiens (man)
C>Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 16-Jul-1999
C:Accession: A60764
R:Huck, S.; Lefranc, G.; Lefranc, M.P.
Immunogenetics 30, 250-257, 1989
A>Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 conve
A:Accession: A60764
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-377 <HUC>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match          50.9%; Score 1144; DB 2; Length 377;
Best Local Similarity 92.5%; Pred. No. 6.8e-66;
Matches 210; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

Oy  2 DKHTCCPCAPPELLGGPSVLFPPPKDPTLMISRTPEVTCTVVVDVSHEDPVEVKFNMYVD 61

```

Db 151 DTPPCPCPAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYD 210  
 Qy 62 GVEVHNATKKPREOYNSTYRVVSVLTVTHQDMLNGKEYCKVSKNALKALPAPIETKTSKAK 121  
 Db 211 GVEVHNATKKPREOYNSTYRVVSVLTVTHQDMLNGKEYCKVSKNALKALPAPIETKTSKTK 270  
 Qy 122 GQPREPQVYVTPPSPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDLDS 181  
 Db 271 GQPREPQVYVTPPSPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDLDS 330  
 Qy 182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNHYTQKSLSLSPGK 228  
 Db 331 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNHYTQKSLSLSPGK 377

RESULT 8  
 G2HU  
 Ig gamma-2 chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Apr-1981 #sequence\_revision 13-Jun-1983 #text\_change 21-Jul-2000  
 C:Accession: A93906; A92809; A90752; A93132; A02148  
 R:Ellison, J.; Hood, L.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 1984-1988, 1982  
 A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain con  
 A:Reference number: A93906; MUID:82197621  
 A:Accession: A93906  
 A:Molecule type: DNA  
 A:Residues: 1-326 <ELL>  
 A:Cross-references: GB:V00554; GB:J00230; NID:g32759; PIDN:CAB58438.1; PID:g6066056  
 A:Note: Lys-326 is probably removed posttranslationally  
 R:Wang, A.C.; Tung, E.; Fudenberg, H.H.  
 J. Immunol. 125, 1048-1054, 1980  
 A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and f  
 A:Reference number: A92809; MUID:81007873  
 A:Contents: myeloma protein T11  
 A:Accession: A92809  
 A:Molecule type: protein  
 A:Residues: 1-19,'Q',21-57,'Z',59,'A',61-193,'D',195-325 <WAN>  
 A:Note: Trp-156 is at or near the complement-binding site  
 R:Connell, G.E.; Parr, D.M.; Hofmann, T.  
 Can. J. Biochem. 57, 758-767, 1979  
 A:Title: The amino acid sequences of the three heavy chain constant region domains of a  
 A:Reference number: A90752; MUID:80001357  
 A:Contents: myeloma protein Zie  
 A:Accession: A90752  
 A:Molecule type: protein  
 A:Residues: 1-24,'E',26-57,'EV',60-85;132-171,'zzz',175,'B',177-193,'D',195-196,'Q',198-  
 A:Note: this sequence has since been revised  
 R:Hofmann, T.; Parr, D.M.  
 Mol. Immunol. 16, 923-925, 1979  
 A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin g  
 A:Reference number: A93132; MUID:80114419  
 A:Contents: Zie  
 A:Accession: A93132  
 A:Molecule type: protein  
 A:Residues: 238-275 <HOF>  
 R:Hofmann, T.; Parr, D.M.  
 submitted to the Atlas, March 1980  
 A:Reference number: A94591  
 A:Contents: annotation; Zie, revisions to residues 25, 59, 60, and 264-268  
 A:Note: the revised sequence differs from that shown in having 60-Ala and in the amidati  
 ned  
 R:Milstein, C.; Frangione, B.  
 Biochem. J. 121, 217-225, 1971  
 A:Title: Disulfide bridges of the heavy chain of human immunoglobulin G2.  
 A:Reference number: A90253; MUID:72033500  
 A:Contents: annotation; myeloma protein Sa, disulfide bonds  
 R:Frangione, B.; Milstein, C.; Pink, J.R.L.  
 Nature 221, 145-148, 1969  
 A:Title: Structural studies of immunoglobulin G.  
 A:Reference number: A93157; MUID:69064124  
 A:Contents: annotation; Sa, disulfide bonds

C:Genetics:  
 A:Gene: GDB:IGHG2  
 A:Cross-references: GDB:119338; OMIM:147110  
 A:Map position: 14q32.33-14q32.33  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light ( chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IM1>  
 F:133-202/Domain: immunoglobulin homology <IM2>  
 F:239-306/Domain: immunoglobulin homology <IM3>  
 F:14/Disulfide bonds: interchain (to light chain) #status experimental  
 F:27-83,140-200,246-304/Disulfide bonds: #status experimental  
 F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental  
 F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.9%; Score 1142.5; DB 1; Length 326;  
 Best Local Similarity 94.1%; Pred. No. 7.2e-66;  
 Matches 209; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

Qy 7 CPCPCAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVGVH 66  
 Db 106 CPCPCAPP-VAGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVQFKWYVGVH 164  
 Qy 67 NAKTKPREEQNSTYRVVSVLTVTHQDMLNGKEYCKVSKNALKALPAPIETKTSKAKQPRE 126  
 Db 165 NAKTKPREEQNSTYRVVSVLTVTHQDMLNGKEYCKVSKNALKALPAPIETKTSKAKQPRE 224  
 Qy 127 PQVYTLPPSDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFF 186  
 Db 225 PQVYTLPPSDEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFF 284  
 Qy 187 LYSKLTVDKSRWQQGNVSCVMHEALHNHYTQKSLSLSPGK 228  
 Db 285 LYSKLTVDKSRWQQGNVSCVMHEALHNHYTQKSLSLSPGK 326

RESULT 9  
 G4HU  
 Ig gamma-4 chain C region - human  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Apr-1982 #sequence\_revision 02-Apr-1982 #text\_change 16-Jul-1999  
 C:Accession: A90933; A90249; A02150  
 R:Ellison, J.; Buxbaum, J.; Hood, L.  
 DNA 1, 11-18, 1981  
 A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.  
 A:Reference number: A90933; MUID:83157104  
 A:Accession: A90933  
 A:Molecule type: DNA  
 A:Residues: 1-327 <ELL>  
 A:Note: the sequence was determined from the germline gene  
 R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.  
 Biochem. J. 117, 33-47, 1970  
 A:Title: Human immunoglobulin subclases. Partial amino acid sequence of the constan  
 A:Reference number: A90249; MUID:70207560  
 A:Accession: A90249  
 A:Molecule type: protein  
 A:Residues: 1-30;81-326 <PIN>  
 C:Genetics:  
 A:Gene: GDB:IGHG4  
 A:Cross-references: GDB:119340; OMIM:147130  
 A:Map position: 14q32.33-14q32.33  
 A:Introns: 99/1; 111/1; 221/1  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light ( chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:20-85/Domain: immunoglobulin homology <IM1>  
 F:99-110/Region: hinge  
 F:134-203/Domain: immunoglobulin homology <IM2>  
 F:240-307/Domain: immunoglobulin homology <IM3>  
 F:14/Disulfide bonds: interchain (to light chain) #status experimental

F;27-93,141-201,247-305/Disulfide bonds: #status predicted  
F;106,109/Disulfide bonds: interchange (to heavy chain) #status experimental  
F;177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 50.5%; Score 1135; DB 1; Length 327;  
Best Local Similarity 93.7%; Pred. No. 2.2e-65;  
Matches 208; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 7 CPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 66  
DB 106 CPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 165

QY 67 NAKTKPREQYNSTYRVSVLTVLDHQLNGKEYCKVSNKALPAPIETKISKAKQPRE 136  
DB 166 NAKTKPREQYNSTYRVSVLTVLDHQLNGKEYCKVSNKALPAPIETKISKAKQPRE 225

QY 127 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSF 186  
DB 226 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSF 285

QY 187 LYSKLVTDKSRWQGNVFCSSVMHEALHNHYTOKSLSLSPGK 228  
DB 286 LYSKLVTDKSRWQGNVFCSSVMHEALHNHYTOKSLSLSLGK 327

## RESULT 10

G3HUWI  
I3 gamma-3 heavy chain disease proteins - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1979 #sequence\_revision 23-Oct-1981 #text\_change 16-Jul-1999  
C;Accession: A90442; A92219; A90198; A93915; A02149  
R;Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.  
Biochemistry 19, 4304-4308, 1980  
A;Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy chain  
A;Reference number: A90442; MUID:81021548  
A;Contents: heavy chain disease protein Wis  
A;Accession: A90442  
A;Molecule type: protein  
A;Residues: 1-289 <PRA>  
A;Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchange  
A;Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co  
R;Michaelsen, T.E.; Frangione, B.; Franklin, E.C.  
J. Biol. Chem. 252, 883-889, 1977  
A;Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication  
A;Reference number: A92219; MUID:77118561  
A;Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W  
A;Accession: A92219  
A;Molecule type: protein  
A;Residues: 12-97 <MIC>  
A;Note: the hinge region in gamma-3 chains is about four times as long as in other gamma  
A;Note: cysteines at positions 24, 27, 33, 39, 42, 48, 54, 57, 63, 69, and 72 form inter  
R;Wolfenstein-Todel, C.; Frangione, B.; Prelli, F.; Franklin, E.C.  
Biochem. Biophys. Res. Commun. 71, 907-914, 1976  
A;Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the  
A;Reference number: A90198; MUID:77021516  
A;Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues  
A;Accession: A90198  
A;Molecule type: protein  
A;Residues: 59-125, 'EB', 128-226, 228-289 <WOL>  
A;Note: this protein lacks most of the V region, all of the CH1 region, and part of the  
R;Alexander, A.; Steinmetz, M.; Barrittault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;  
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982  
A;Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion  
A;Reference number: A93915; MUID:82247835  
A;Contents: heavy chain disease protein Omn  
A;Accession: A93915  
A;Molecule type: mRNA  
A;Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157  
A;Note: a carboxyl-terminal Lys is removed posttranslationally  
A;Note: this sequence may represent an allelic form or another gamma chain subclass

C;Comment: The heavy chain disease protein Wis is shown.

C;Genetics:  
A;Gene: GDB:IGHG3  
A;Cross-references: GDB:119339; OMIM:147120  
A;Map position: 14q32.33-14q32.33  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid  
F;203-270/Domain: immunoglobulin homology <IMM>  
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F;6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 49.9%; Score 1121; DB 1; Length 289;  
Best Local Similarity 90.3%; Pred. No. 1.5e-64;  
Matches 204; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

QY 2 DKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
DB 64 DTPPPCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 123

QY 62 GVEVHNKTKPREQYNSTYRVSVLTVLDHQLNGKEYCKVSNKALPAPIETKISKAK 121  
DB 124 GVQVHNKTKPREQYNSTYRVSVLTVLDHQLNGKEYCKVSNKALPAPIETKISKAK 183

QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181  
DB 184 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 243

QY 182 DGSFFLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTOKSLSLSPG 227  
DB 244 DGSFFLYSKLTVDKSRWQGNVFCSSVMHEALHNHYTOKSLSLSPG 289

## RESULT 11

GHRB  
Ig gamma chain C region - rabbit  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 24-Apr-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Jul-1999  
C;Accession: A91749; A90290; A93928; A90245; A94416; A02161  
R;Bernstein, K.E.; Alexander, C.B.; Mage, R.G.  
Immunogenetics 18, 387-397, 1983  
A;Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I hap  
A;Reference number: A91749; MUID:84030930  
A;Accession: A91749  
A;Molecule type: mRNA  
A;Residues: 1-323 <BER>  
A;Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185-  
R;Pratt, D.M.; Mole, L.E.  
Biochem. J. 151, 337-349, 1975  
A;Title: Sequence studies on the constant region of the Fd sections of rabbit immunog  
A;Reference number: A90290; MUID:76135459  
A;Accession: A90290  
A;Molecule type: protein  
A;Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>  
R;Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.  
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982  
A;Title: Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma heavy ch  
A;Reference number: A93928; MUID:83299917  
A;Accession: A93928  
A;Molecule type: mRNA  
A;Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>  
A;Cross-references: GB:M16426; NID:gl65111; PIDN:AAA31289, 1; PID:gl65112  
A;Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic ma  
R;Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.  
Biochem. J. 116, 249-259, 1970  
A;Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobul  
A;Reference number: A90245; MUID:70110015  
A;Accession: A90245  
A;Molecule type: protein  
A;Residues: 132-143, 'E', 145-161 <FRU>  
R;Hill, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.  
In Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wikse  
A;Reference number: A94416

[illegible]

Db 43 VDRKRVTKTKPPCPICPACGPGPSAFIPFPKPKDTLMISRTKPKVTCVVVDVDSQENPEVQ 102  
QY 56 FNNYVDGVEVHNAKTPREQYNSYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 115  
Db 103 FSNYVDGVEVHTAQTTPKEQFNSTRVSVLPIQHDWLNKGEFKCKVNNKDLPAPIR 162  
QY 116 TISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQ--PENNYK 173  
Db 163 IISKAKGQTPREQVYTLPPSTEELSKVTLTCLVTGTFYPPDIDVEMQNGQPEPEGNR 222  
QY 174 TTPPVLDSDGSPFLYSLKLTVDKSRWQGNVFCVSMHEALHNYHTOKSLSPGK 228  
Db 223 TTPPQDDVDGTYFLYSLKLAVDKASWQGRDTFQCAVMHEALHNYHTOKSIFKTPGK 277

RESULT 15  
G2QP  
Ig gamma-2 chain C region - guinea pig  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 07-May-1981 #sequence\_revision 07-May-1981 #text\_change 16-Jul-1999  
C:Accession: A94553; A90352; A90359; A90384; A90385; A02151  
R:Trischmann, T.M.  
submitted to the Atlas, April 1975  
A:Reference number: A94553  
A:Accession: A94553  
A:Molecule type: protein  
A:Residues: 1-3 <TRI>  
R:Brishstein, B.K.; Hussain, O.Z.; Cebra, J.J.  
Biochemistry 10, 18-25, 1971  
A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. Am  
A:Reference number: A90352; MUID:71058471  
A:Accession: A90352  
A:Molecule type: protein  
A:Residues: 4-68 <BIR>  
R:Turner, K.J.; Cebra, J.J.  
Biochemistry 10, 9-17, 1971  
A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II. Am  
A:Reference number: A90359; MUID:71058486  
A:Accession: A90359  
A:Molecule type: protein  
A:Residues: 69-133;312-329 <TUR>  
R:Tracey, D.E.; Cebra, J.J.  
Biochemistry 13, 4796-4803, 1974  
A:Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibodies.  
A:Reference number: A90384; MUID:75036072  
A:Accession: A90384  
A:Molecule type: protein  
A:Residues: 134-226 <TRA>  
R:Trischmann, T.M.; Cebra, J.J.  
Biochemistry 13, 4804-4811, 1974  
A:Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibodies.  
A:Reference number: A90385; MUID:75036073  
A:Accession: A90385  
A:Molecule type: protein  
A:Residues: 227-311 <TR2>  
R:Oliveira, B.; Lamm, M.E.  
Biochemistry 10, 26-31, 1971  
A:Title: Interchain disulfide bridges of guinea pig gamma-2- immunoglobulin.  
A:Reference number: A90354; MUID:71058474  
A:Contents: annotation; disulfide bonds  
A:Note: Cys-16 is involved in a heavy-light chain bond  
A:Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds  
C:Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.  
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger complexes.  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:21-81/Domain: immunoglobulin homology <IM1>  
F:135-204/Domain: immunoglobulin homology <IM2>  
F:241-310/Domain: immunoglobulin homology <IM3>  
F:28-79/Disulfide bonds: #status experimental  
F:142-202/Disulfide bonds: #status experimental  
F:178/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:248-308/Disulfide bonds: #status experimental

Query Match 39.6%; Score 889; DB 1; Length 329;  
Best Local Similarity 72.3%; Pred. No. 1e-49;  
Matches 162; Conservative 24; Mismatches 36; Indels 2; Gaps 1;  
QY 6 TCPPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEV 65  
Db 106 TCPPCPAPPENLGGPSVFIFFPKPKDTLMISLTFRVTCVVVDVSDQDEPEVQTFWFDNKPV 165  
QY 66 HNAKTPREEQYNSYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPR 125  
Db 166 GNAETKPRVEQYNTTFRVESVLPFIHQDWLGRKEFKCKVYNKALPAPIEKTIISKAGAPR 225  
QY 126 EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQP--ENNYKTTTPPVLDSDG 183  
Db 226 MPDVYTLPPSRDELTKSKSVTCLIIINFFPADIHVEWASNRVPSRSEKEYKNTPTPIEDADG 285  
QY 184 SFELYSKLTVDKSRWQGNVFCVSMHEALHNYHTOKSLSPG 227  
Db 286 SYFLYSLKLTVDKSAWDQGITVTCVMHEALHNYHTOKAISRSPG 329

Search completed: March 1, 2001, 09:15:50  
Job time: 148 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 09:17:08 ; Search time 70.93 Seconds  
(without alignments)  
193.313 Million cell updates/sec

Title: US-09-389-782A-8

Perfect score: 2246

Sequence: 1 MDKTHCTPCPCAPPELLGPPS.....QKGNATHDNCISGNSESTQK 401

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_36.\*  
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1239.5	55.2	652	19 W48650	Heavy chain of hma
2	1238	55.1	374	19 W83963	Recombinant human
3	1238	55.1	374	19 W49075	Recombinant human
4	1237	55.1	235	20 Y01372	Amino acid sequenc
5	1236.5	55.1	651	18 W26649	Chimeric receptor
6	1236.5	55.1	692	18 W26650	Chimeric receptor
7	1234	54.9	347	13 R27163	CD2 binding LFA-3-
8	1234	54.9	347	21 Y83136	Human transmembran
9	1234	54.9	388	19 W73513	Rabbit TGFbetarII:
10	1234	54.9	388	19 W73514	Human TGFbetarII:F
11	1234	54.9	388	21 Y54063	Amino acid sequenc
12	1234	54.9	388	21 Y54064	Amino acid sequenc

13	1234	54.9	399	21 Y70867	Human interferon-b
14	1234	54.9	418	21 Y70868	Human interferon-b
15	1234	54.9	423	21 Y70869	Human interferon-b
16	1234	54.9	446	15 R58753	VCAM 2D-IgG. Homo
17	1234	54.9	446	20 Y23986	VCAM 2D-IgG, a sol
18	1234	54.9	446	20 Y01037	VCAM 2D-IgG protei
19	1234	54.9	446	20 W96743	A VCAM 2D-IgG1 fus
20	1234	54.9	482	19 W31646	Human cytokine rec
21	1233	54.9	232	18 W26232	Human IgG1 hinge/F
22	1233	54.9	233	20 Y06617	Human Fc (IgG1).
23	1233	54.9	259	20 Y24154	Protein from pCD51
24	1233	54.9	329	17 R91806	Human immunoglobul
25	1233	54.9	351	14 R43685	Human kappa immuno
26	1233	54.9	371	10 P91918	Sequence of the li
27	1233	54.9	371	10 P93558	Linkered human IgG
28	1233	54.9	376	19 W60037	Antigenic peptide
29	1233	54.9	379	19 W83962	Recombinant human
30	1233	54.9	379	19 W49073	Recombinant human
31	1233	54.9	379	19 W49074	Recombinant human
32	1233	54.9	387	17 R90920	IL4_Y124D/IgG1 pro
33	1233	54.9	396	18 W18574	Aggrecanase artifi
34	1233	54.9	396	18 W18575	Aggrecanase artifi
35	1233	54.9	400	21 Y15123	Porcine CTLA-4-Ig
36	1233	54.9	424	16 W14764	Human soluble kit
37	1233	54.9	424	16 W14765	Human soluble kit
38	1233	54.9	435	13 R26530	Sequence of one ch
39	1233	54.9	437	18 W10552	Alpha-1-acid glyco
40	1233	54.9	442	18 W10550	IgG1 polypeptide.
41	1233	54.9	445	20 Y24153	Bovine LOX-1 extra
42	1233	54.9	446	17 W05829	Humanised LD10 ant
43	1233	54.9	447	20 Y31669	Human IgG1 chain C
44	1233	54.9	449	14 R43339	Completely humanis
45	1233	54.9	449	19 W49816	Amino acid sequenc

## ALIGNMENTS

RESULT 1	
W48650	
ID W48650 standard; Protein; 652 AA.	
XX	
AC W48650;	
XX	
DT 04-AUG-1998 (first entry)	
XX	
DE Heavy chain of hmaB425 fused to TNF alpha.	
XX	
KW Antibody-cytokine fusion protein; triclstronic vector; chimeric;	
KW TNF alpha; IL-2; IRES; internal ribosome entry site.	
XX	
OS Synthetic.	
OS Homo sapiens.	
XX	
FH Key	Location/Qualifiers
FT Region	1..494
FT Region	/note= "Heavy chain of human mab 425"
FT Region	495..652
FT Region	/note= "TNF alpha"
XX	
PN WO9811241-A1.	
XX	
PD 19-MAR-1998.	
XX	
PF 02-SEP-1997;	97WO-EP04765.
XX	
PR 30-SEP-1996;	96EP-0115635.
PR 16-SEP-1996;	96EP-0114820.
XX	
PA (MERE ) MERCK PATENT GMBH.	
XX	
PI Bruemmer W, Burge C, Dunker R,	
PI Rieke E, Von Hoegen I, Welge T;	
	Hauser H, Mielke C;

XX WPI; 1998-207400/18.  
 DR N-PSDB; V18096.  
 XX  
 XX Oligo:cistronic expression vector - useful for production of, e.g.  
 PT MAB425/TNF- $\alpha$  or MAB425/IL-2 antibody fusion protein  
 PT  
 XX Disclosure; Fig 15; 89pp; English.  
 PS  
 XX The present sequence represents a fusion protein comprising of TNF  
 CC alpha fused to the C-terminus of the heavy chain of the human  
 CC monoclonal antibody 425 (hMAB425). The hMAB425 has specificity for  
 CC the human EGF receptor. The invention claims for a new pMCDHAP  
 CC tricistronic vector (V18096) for the expression of an antibody-cytokine  
 CC fusion protein, hMAB425-TNF alpha. The TNF alpha sequence can be  
 CC substituted by the IL-2 sequence. The vector also contains a strong  
 CC promoter/enhancer unit, a selection marker gene and at least two  
 CC poliovirus derived internal ribosomal entry site (IRES) sequences. The  
 CC vector can be expressed in mammalian host cells for the production of  
 CC heteromeric fusion proteins. This expression system is claimed to  
 CC produce the heteromeric proteins in high yields.  
 XX  
 SQ Sequence 652 AA;

Query Match 55.2%; Score 1239.5; DB 19; Length 652;  
 Best Local Similarity 65.8%; Pred. No. 9.4e-76;  
 Matches 250; Conservative 19; Mismatches 54; Indels 57; Gaps 6;  
 QY 2 DKHTCPCPAPPELLGGPSVFLFPPPKDITLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61  
 Db 268 dkhtcpcpapelggpsvflfpppkdtlmisrtpevtcvvvdshedpevkfnwyvd 327  
 QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121  
 Db 328 gvevhnaaktpreeqynstyrvvsvltvlhqdwlngkeyckvsnkalpapiektiskak 387  
 QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLIVKGFPYSDIAVEWESNGQPENNYKTPPPVLD 181  
 Db 388 gqprepqvtytlppsrdeitknqvslclivkgfypsdiavewesngqpennyktppvlds 447  
 QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKSLSPGKETFPFKYLHYDEE 241  
 Db 448 dgsfflyskltvdksrwqgnvfscvsmhealhnhytkslspgketrpfpkylhydee 507  
 QY 242 TSHQLLCKDKPPGYTLKQHTAKWKTCAPCPDHYITDTSWHTSDCLYCSVCK-ELQYV 300  
 Db 508 vah-----vvanpqaegqlwl 524  
 QY 301 KOECNR-----THNRVCECKEGRYL-ETEFCLKHRSCPPGGVVOAGTPTERTVCKR 351  
 Db 525 nrranallangvelrdnqlvvpseglylysqvlfkgggc-psthlvllthtsrlavsyq 583  
 QY 352 CPDGFPSNETSSKAPCRKHT 371  
 Db 584 tkvnlls---aikspcqrct 600  
 RESULT 2  
 W83963  
 ID W83963 standard; Protein; 374 AA.  
 XX  
 AC W83963;  
 XX  
 DT 05-FEB-1999 (first entry)  
 XX  
 DE Recombinant human metFc-OB protein variant.  
 XX  
 KW Recombinant; metFc-OB protein; Fc region; immunoglobulin; Ig; OB;  
 KW obesity; human; adiposity; blood lipid; diabetes type II; insulin;  
 KW hypoglycaemic; antihypertensive; diuretic; appetite suppressant;  
 KW suspension; variant.  
 XX

OS Homo sapiens.  
 XX  
 PN W09846257-A1.  
 XX  
 PD 22-OCT-1998.  
 XX  
 PF 16-APR-1998; 98WO-US07828.  
 XX  
 PR 14-APR-1997; 98US-0059467.  
 PR 17-APR-1997; 97US-0843971.  
 XX  
 PA (AMGE-) AMGEN INC.  
 XX  
 PI Brems DN, French DL, Speed MA;  
 DR WPI; 1998-594525/50.  
 XX N-PSDB; V69686.  
 PT Concentrated suspension of fusion of obesity protein with Fc  
 PT immunoglobulin fragment - stable at physiological pH, used for e.g.  
 PT reduction of weight and blood lipid levels, and for treatment of  
 PT type II diabetes  
 XX  
 PS Claim 2; Fig 6A-C; 47pp; English.  
 XX  
 CC This represents a recombinant metFc-OB protein variant which consists of  
 CC an Fc region of human immunoglobulin (Ig) fused to a human OB (obesity)  
 CC protein. The invention provides a human OB protein suspension that  
 CC contains at least 0.5 mg/ml of the human OB protein derivatised by  
 CC attachment of the Fc region of an Ig to the N-terminus of OB, and has a  
 CC pH 6-8. The suspensions are used to reduce weight, adiposity and blood  
 CC lipid levels, to treat or prevent diabetes type II, and to increase lean  
 CC mass and insulin sensitivity. They may be used in conjunction with  
 CC insulin, hypoglycaemics, antihypertensives, diuretics, appetite  
 CC suppressants etc. These suspensions are stable and active at  
 CC physiological pH and are ready-for-use formulations that do not require  
 CC freezing or freeze drying. As they are very concentrated, only small  
 CC volumes are required and they provide a sustained-release effect, with  
 CC increased potency and reduced frequency of injection.  
 XX  
 SQ Sequence 374 AA;  
 Query Match 55.1%; Score 1238; DB 19; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-76;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDKTHTCPCPAPPELLGGPSVFLFPPPKDITLMISRTPEVTCVVDVSHEDPEVKFNWYV 60  
 Db 1 mdkthtcpcpapelggpsvflfpppkdtlmisrtpevtcvvvdshedpevkfnwyv 60  
 QY 61 DGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 120  
 Db 61 dgvevhnaaktpreeqynstyrvvsvltvlhqdwlngkeyckvsnkalpapiektiska 120  
 QY 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLIVKGFPYSDIAVEWESNGQPENNYKTPPPVLD 180  
 Db 121 kgqprepqvtytlppsrdeitknqvslclivkgfypsdiavewesngqpennyktppvld 180  
 QY 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNNHYTKSLSPGK 228  
 Db 181 sdgsfflyskltvdksrwqgnvfscvsmhealhnhytkslspgk 228  
 RESULT 3  
 W49075  
 ID W49075 standard; Protein; 374 AA.  
 XX  
 AC W49075;  
 XX  
 DT 18-NOV-1998 (first entry)  
 XX  
 DE Recombinant human MetFc-OB variant 2 protein.



XX Recombinant human MetFc-OB variant 2 protein; chimeric; immunoglobulin;  
 KW high blood lipid level; arterial sclerosis; stroke; Fc-OB fusion protein;  
 KW diabetes.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 XX Location/Qualifiers  
 FT Protein  
 FT 2..374  
 FT /note= "Recombinant human Fc-OB variant 2 protein"  
 FT Region  
 FT 229..374  
 FT /note= "Human OB protein"  
 XX  
 XX WO9828427-A1.  
 XX  
 XX 02-JUL-1998.  
 XX  
 XX 11-DEC-1997; 97WO-US23183.  
 XX  
 XX 20-DEC-1996; 96US-0770973.  
 XX  
 XX (AMGE-) AMGEN INC.  
 XX  
 XX Hecht RI, Mann MB;  
 XX  
 XX WPI; 1998-377658/32.  
 XX N-PSDB; V32902.  
 XX  
 XX New fusion proteins of OB and Fc - used for treating e.g. excess  
 PT weight, diabetes, arterial sclerosis, arterial plaque, high blood  
 PT lipid level, gall stones or stroke  
 XX  
 XX Claim 2; Fig 5A-5C; 107pp; English.  
 XX  
 CC The present sequence represents a recombinant human MetFc-OB variant 2  
 CC fusion protein having a 5 residue deletion of residues 2-6 of the  
 CC wild-type Fc-OB protein sequence shown in W49073. The invention provides  
 CC Fc-OB fusion proteins whereby the Fc region of an immunoglobulin or its  
 CC analogue is linked, either directly or indirectly using a linker, to the  
 CC N-terminus of an OB protein or its analogue. The Fc-OB fusion proteins  
 CC are claimed to demonstrate increased stability and clearance rate and  
 CC decreased degradation as compared to OB protein or a fusion of Fc to  
 CC the C-terminus of the OB protein. These Fc-OB fusion proteins are also  
 CC claimed to be useful for treating excess weight in an individual or  
 CC animal or for treating co-morbidities associated with excess fat such as  
 CC diabetes, high blood lipid level, arterial sclerosis and stroke.  
 XX  
 SQ Sequence 374 AA;

Query Match 55.1%; Score 1238; DB 19; Length 374;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-76;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MDKTHTCPCPAPALGGPSVFLPPPKDMLSRPEVTCVVVDSHEDPEVKFNWYV 60  
 DB 1 mdkthtcpcpapalggpsvflfppkdtlmsrptevtcvvvdsheadpevkfnwv 60  
 QY 61 DGEVHNKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKA 120  
 DB 61 dgevhnaktpreeqnystyrvvsvltvlhqdwlngkeyckvsnkalpapiektiska 120  
 QY 121 KGQPREQVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 180  
 DB 121 kgqpreqvyltlpsrdeltnqvscltclvkgyfypsdiavwesngqpennnykttppvld 180  
 QY 181 SDGSFFLYSKLTVDKSRWQQGNVFCSMVHREALNHNHTQKLSLSLSPGK 228  
 DB 181 sdgsfflyskltvdksrwqqgnvfscsmvhrealnhnhtqklslsispgk 228  
 RESULT 4

Y01372  
 ID Y01372 standard; Protein; 235 AA.  
 XX  
 AC Y01372;  
 XX  
 DT 04-JUN-1999 (first entry)  
 XX  
 DE Amino acid sequence of Fc fragment of human IgG.  
 XX  
 XX FeRn; binding; epithelial cell; immune system; modulation; antigen;  
 KW pathogen; autoimmune disease; allergen; tumour; therapeutic; cytokine;  
 KW chemotherapy agent; interferon; insulin; human growth hormone; fertility;  
 KW drug; calcitonin; steroid; immunity; mucosal; AIDS; hepatitis; Fc; human;  
 KW immunoglobulin; IgG.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 230  
 FT /note= "encoded by TGA"  
 XX  
 XX WO9904813-A1.  
 XX  
 XX 04-FEB-1999.  
 XX  
 XX 24-JUL-1998; 98WO-US15395.  
 XX  
 XX 24-JUL-1997; 97US-0899856.  
 XX  
 XX (BGHM ) BRIGHAM & WOMENS HOSPITAL.  
 PA (UYBR-) UNIV BRANDEIS.  
 XX  
 PI Blumgerg RS, Lencer WI, Simister NE;  
 XX  
 XX WPI; 1999-153297/13.  
 XX N-PSDB; X27800.  
 XX  
 XX Delivering therapeutics, particularly antigens to epithelial cells -  
 PT comprises use of a conjugate of the therapeutic and an FcRn binding  
 PT partner  
 XX  
 XX Disclosure; Fig 1; 74pp; English.  
 XX  
 CC The invention relates to methods of delivering molecules to a mammal by  
 CC administering a conjugate of a therapeutic or a bioactive substance,  
 CC and an FcRn binding partner targeted to epithelial cells expressing FcRn.  
 CC This can be used for modulating a mammalian immune system wherein, the  
 CC bioactive substance is an antigen characteristic of a pathogen of an  
 CC autoimmune disease, an allergen, or a tumour. The delivering method is  
 CC used to deliver therapeutics, including chemotherapy agents, cytokines  
 CC including interferon, hormones, including insulin and human growth  
 CC hormone, fertility drugs, calcitonin, calcitriol and other bioactive  
 CC steroids to intestinal, mucosal and lung epithelium. The method is also  
 CC used for trans epithelial delivery of antigens to provoke tolerance and  
 CC immunity. The method provides an immunisation that specifically targets  
 CC the mucosal surfaces, and does not have the risks associated with  
 CC injections, including needle transmission of AIDS or hepatitis, of prior  
 CC art immunisation methods. The present sequence represents the amino acid  
 CC sequence of an Fc fragment of human immunoglobulin IgG.  
 XX  
 SQ Sequence 235 AA;

Query Match 55.1%; Score 1237; DB 20; Length 235;  
 Best Local Similarity 97.4%; Pred. No. 4.1e-76;  
 Matches 228; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 MDKTHTCPCPAPALGGPSVFLPPPKDMLSRPEVTCVVVDSHEDPEVKFNWYV 60  
 DB 2 vdkthtcpcpapalggpsvflfppkdtlmsrptevtcvvvdsheadpevkfnwv 61  
 QY 61 DGEVHNKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKA 120  
 DB 61 dgevhnaktpreeqnystyrvvsvltvlhqdwlngkeyckvsnkalpapiektiska 120

Db 62 dgvevhnaktkpreeqnystyrsvsvltvlhqdwlngkeyckvsnkalpapietkiska 121  
 QY 121 KGPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWESNGQPNENYKTTTPPVLD 180  
 Db 122 kgprepqvtytlppsrdeltnqvsitclvkgfypsdiawesngqpennkyttppvld 181  
 QY 181 SDGSFFLYSLKTVDKSRWQGVNFCVSMHEALHNYTKSLSPGKETPPPK 234  
 Db 182 sdgsfflyskltvdkswqgvnfcsvmhealhnytkslspgkxvrpr 235

RESULT 5  
 W26649  
 ID W26649 standard; Protein; 651 AA.

AC W26649;  
 XX  
 DT 12-FEB-1998 (first entry)  
 DE Chimeric receptor hCTM01/G1/zeta.

Cell activation: chimeric receptor; DNA delivery; CTM01; scFv;  
 IgG1; T cell receptor zeta chain; cell proliferation; cytokine;  
 inflammation; effector; cell differentiation; antibody secretion;  
 phagocytosis; tumour infiltration; adhesion; infection; cancer;  
 allergy; rheumatoid arthritis; osteoarthritis; asthma; eczema;  
 inflammatory bowel disease; cystic fibrosis; sickle cell anaemia;  
 psoriasis; multiple sclerosis; transplant rejection; diabetes;  
 graft versus host disease; human; therapy.

Chimeric - Homo sapiens.

WO9723613-A2.

03-JUL-1997.

23-DEC-1996; 96WO-GB03209.

21-DEC-1995; 95GB-0026131.

(CLLT ) CELLTech THERAPEUTICS LTD.

Bebington CR, Finney HM, Lawson ADG, Weir ANC;

WPI; 1997-351052/32.

N-PSDB; T90512.

New DNA systems for activating cells - comprising DNA coding for a  
 chimeric receptor comprising 2 or more different cytoplasmic  
 signalling components.

Disclosure; Fig 7; 90pp; English.

This protein comprises a chimeric receptor consisting of an scFv  
 engineered from anti-CD3 human antibody CTM01 linked to an  
 extracellular spacer comprising the human IgG1 hinge, CH2 and CH3,  
 linked to transmembrane and intracellular regions of the human T  
 cell receptor zeta chain. It can be expressed in host cells (e.g.  
 Jurkat) using a chimeric receptor gene (see T90512) constructed  
 from DNA cassettes encoding each receptor component. In a claimed  
 cell activation process an effector cell is transformed with DNA  
 encoding a chimeric receptor containing 2 or more different  
 cytoplasmic signalling components. Also claimed is use of DNA  
 encoding a recombinant chimeric receptor in a DNA delivery system.  
 The DNA delivery systems can be used for the activation of cells to  
 provide e.g. an increase in cell proliferation, expression of  
 cytokines with e.g. pro- or anti-inflammatory responses, stimulation  
 of cytolytic activity, differentiation or other effector functions,  
 antibody secretion, phagocytosis, tumour infiltration and/or  
 increased adhesion. They can be used in the treatment of e.g.  
 infectious disease, inflammatory disease, cancer, allergic/atopic  
 disease, congenital disease, dermatologic disease, neurologic  
 disease, transplants and metabolic/idiopathic disease (claimed).

CC In particular, they can be used in the treatment of rheumatoid  
 CC arthritis, osteoarthritis, inflammatory bowel disease, asthma,  
 CC eczema, cystic fibrosis, sickle cell anaemia, psoriasis, multiple  
 CC sclerosis, organ or tissue transplant rejection, graft-versus-host  
 CC disease or diabetes (claimed).

XX Sequence 651 AA;

Query Match 55.1%; Score 1236.5; DB 18; Length 651;  
 Best Local Similarity 62.6%; Pred. No. 1.5e-75;  
 Matches 258; Conservative 18; Mismatches 61; Indels 75; Gaps 9;

QY 2 DKTHCPCPAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDSHEDPCKVNWYVD 61  
 Db 288 dkthcpcppapellgspvflfpkpkdtlmisrtpevtcvvdsvedpckvfnwyvd 347

QY 62 GVEVHNAKTPREEQYNSTYRVSVSLTVLHODWLNKGYCKVSNKALPAPIETKISKAK 121  
 Db 348 gvevhnaktkpreeqnystyrsvsvltvlhqdwlngkeyckvsnkalpapietkiskak 407

QY 122 GPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWESNGQPNENYKTTTPPVLD 181  
 Db 408 gprepqvtytlppsrdeemtknqvsitclvkgfypsdiawesngqpennkyttppvlds 467

QY 182 DGSFFLYSLKTVDKSRWQGVNFCVSMHEALHNYTKSLSPGKETPPPKLYLHDEE 241  
 Db 468 dgsfflyskltvdkswqgvnfcsvmhealhnytkslspgk-ldpk----- 518

QY 242 TSHOLLCDKPPGTYLKHQCTAKWKTCVPCPDHYTDSWHTSDCLYCSVPKELQYVK 301  
 Db 519 -----lc-----ylldgilfygviltalfir 540

QY 302 -----QECNPTHNVCECKEGRYLETFCLKHKRSCPPCGFVGWAGTPERTV 348  
 Db 541 vkfsrsadapayqgqnglyn---elnlgrreeydvldkrrgrdpemg----gkpr--- 590

QY 349 CKRCPDGFSSNETSSKAPCRKHTNCVFGLLLTOKGNATHDNCISGNSESTQ 400  
 Db 591 -knpgqgly-nelqdkmaeayseigmgerrrkg---hdglyqglstak 637

RESULT 6

W26650  
 ID W26650 standard; Protein; 692 AA.

XX W26650;

XX 12-FEB-1998 (first entry)

XX Chimeric receptor hCTM01/G1/zeta-CD28.

Cell activation; chimeric receptor; DNA delivery; CTM01; scFv;  
 IgG1; CD28; T cell receptor zeta chain; proliferation; cytokine;  
 inflammation; effector; cell differentiation; antibody secretion;  
 phagocytosis; tumour infiltration; adhesion; infection; cancer;  
 allergy; rheumatoid arthritis; osteoarthritis; asthma; eczema;  
 inflammatory bowel disease; cystic fibrosis; sickle cell anaemia;  
 psoriasis; multiple sclerosis; transplant rejection; diabetes;  
 graft versus host disease; human; therapy.

Chimeric - Homo sapiens.

WO9723613-A2.

03-JUL-1997.

23-DEC-1996; 96WO-GB03209.

21-DEC-1995; 95GB-0026131.

(CLLT ) CELLTech THERAPEUTICS LTD.

PI Bebbington CR, Finney HM, Lawson ADG, Weir AN; WPI; 1997-351052/32. N-PSDB; T90513.

XX New DNA systems for activating cells - comprising DNA coding for a PT chimeric receptor comprising 2 or more different cytoplasmic PT signalling components.

XX Disclosure; Fig 8; 90pp; English.

XX This protein comprises a chimeric receptor consisting of an scFv CC engineered from anti-CD3 human antibody CTMO1 linked to an CC extracellular spacer comprising the Human IgG1 hinge, CH2 and CH3, CC linked to the transmembrane and intracellular components of the CC human T cell receptor zeta chain, fused to the intracellular region CC of human CD28. It can be expressed in host cells (e.g. Jurkat) CC using a chimeric receptor gene (see T90513) constructed from DNA CC cassettes encoding each component of the receptor. In a claimed CC cell activation process an effector cell is transformed with DNA CC encoding a chimeric receptor containing 2 or more different CC cytoplasmic signalling components. Also claimed is use of DNA CC encoding a recombinant chimeric receptor in a DNA delivery system. CC The DNA delivery systems can be used for the activation of cells to CC provide e.g. an increase in cell proliferation, expression of CC cytokines with e.g. pro- or anti-inflammatory responses, stimulation CC of cytolytic activity, differentiation or other effector functions; CC antibody secretion, phagocytosis, tumour infiltration and/or CC increased adhesion. They can be used in the treatment of e.g. CC infectious disease, inflammatory disease, cancer, allergic/atopic CC disease, congenital disease, dermatologic disease, neurologic CC disease, transplants and metabolic/idiopathic disease (claimed). CC In particular, they can be used in the treatment of rheumatoid CC arthritis, osteoarthritis, inflammatory bowel disease, asthma, CC eczema, cystic fibrosis, sickle cell anaemia, psoriasis, multiple CC sclerosis, organ or tissue transplant rejection, graft-versus-host CC disease or diabetes (claimed).

XX SQ Sequence 692 AA;

Query Match 55.1%; Score 1236.5; DB 18; Length 692;  
Best Local Similarity 62.6%; Pred. No. 1.6e-75;  
Matches 258; Conservative 18; Mismatches 61; Indels 75; Gaps 9;

QY 2 DKHTCPCCPAPPELLGGPSVFLPDKPTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61  
Db 288 dkhtcpccpapelggpsvflfpdkptlmisrtpetvctvvdvshedpevkfnwyvd 347

QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVHLQDNLNKEYCKVSNKALPAPIETKISKAK 121  
Db 348 gvevhnaktprreeqnstyrvvsvltvhlqdwlngkeyckvsnkalpapietkiskak 407

QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGOPENNYKTTTPVLD 181  
Db 408 gqprepqvlytlppsrdeemtknqvslclvkgfypsdiavewesngopennykttppvlds 467

QY 182 DGSFELYSKLTVDKSRWQGVNVCSCVWHEALHNHYTQKSLSLSPGKETPPKYLHDEE 241  
Db 468 dgsfelyskltvdksrwqgvnvcscvwmhealhnhytcqkslsispgk--ldpk----- 518

QY 242 TSHQLLCKDPCPGGYLKHQHTAKWTKVCAPCPDHYHYDTSWHTSDCELYCSPVKELQYVK 301  
Db 519 ----lc-----yldgliflygvltalflr 540

QY 302 -----QECNRTNIRVCECKEGRYLEIEFCLKHRSCPPPGFVQVQAGTPERTV 348  
Db 541 vkfsrsadapayqggqqlyn---elnlgreeydvldkrrgrdpeng----gkpr--- 590

QY 349 CKRCPDGFNSWSSKAPCRKHTNCVFGLLLTQKGNATHDNCISGNSSTQ 400  
Db 591 -knpqegly-neiqdkmaeaysigmkerrrgk---hdglyqglstatk 637

RESULT 7  
R27163  
ID R27163 standard; Protein; 347 AA.  
XX  
AC R27163;  
XX  
DT 20-MAY-1998 (first entry)  
XX  
DE CD2 binding LFA-3-Ig fusion protein.  
XX  
KW lymphocyte associated antigen-3; T-lymphocyte accessory molecule;  
KW deletion mutant; CD2 binding site; immunomodulator; immunoglobulin;  
KW preLFA3TIP; LFA-3(92)IgG; pSAB152; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT peptide 1..28 /label= LFA-3\_signal  
FT region 29..120 /label= LFA-3\_aminoacids\_1-92  
FT region 121..130 /label= IgG1\_hinge  
FT domain 131..239 /label= IgG1\_CH2  
FT domain 240..347 /label= IgG1\_CH3  
XX  
PN EP503648-A.  
XX  
PD 16-SEP-1992.  
XX  
PF 12-MAR-1992; 92EP-0104320.  
XX  
PR 12-MAR-1991; 91US-0667971.  
PR 07-OCT-1991; 91US-0770967.  
XX  
PA (BIOJ ) BIOGEN INC.  
XX  
PI Miller GT, Rosa MD, Wallner BP;  
XX  
DR WPI; 1992-309760/38.  
DR N-PSDB; Q28684.  
XX  
PT CD2-binding domain of lymphocyte function associated antigen-3  
PT and DNA - for diagnosing and treating inflammation and  
PT autoimmune diseases, e.g. systemic lupus erythematosus and  
PT rheumatoid arthritis  
XX  
PS Claim 13; Fig 12; 85pp; English.  
XX  
CC The plasmid pSAB152 contains the DNA sequence encoding the LFA-3  
CC signal sequence, the amino terminal 92 amino acids of mature LFA-3,  
CC ten amino acids of the hinge region of IgG1 and the CH2 and CH3  
CC constant domains of IgG1 (see Q28678-9 and Q28681-2 for details of  
CC the construction of pSAB152). A NotI fragment containing the coding  
CC sequence of pSAB152 was used in the construction of expression  
CC vector pMDR(92)ig-3 which can be stably maintained in CHO cells to  
CC achieve continuous expression of LFA3TIP. The fusion protein can  
CC bind to CD2 and inhibit T cell activation, making it useful to  
CC treat acute and chronic inflammation, autoimmune disease and  
CC in immunomodulation.  
XX  
SQ Sequence 347 AA;

Query Match 54.9%; Score 1234; DB 13; Length 347;  
Best Local Similarity 59.6%; Pred. No. 1e-75;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHTCPPCAPPELLGGPSVFLPDKPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
:|||||

Db 120 vdkthtccpapellggpsvflfpkpkdtlmiartpevtcvvvdshedpevkfnwv 179

Qy 61 DGVEVHNKTKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 120  
 |||||  
 Db 180 dgvevhnaktpreeqynstyrsvsvltvlhqdwlngkeyckvsnkalpapiektiska 239  
 |||||

Qy 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 180  
 |||||  
 Db 240 kgqprepqvylppsrdeltnqvsiltclvkgyfypsdiavewesngqpennyktppvld 299  
 |||||

Qy 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228  
 |||||  
 Db 300 sdgsfflyskltdvksrwwqgnvfscsvmhealhnhytqkslsispgk 347  
 |||||

RESULT 8

Y83136 ID Y83136 standard; Protein; 347 AA.

XX AC Y83136;

XX DT 24-JUL-2000 (first entry)

XX DE Human transmembrane LFA-3/IgG fusion protein LFA3TIP.

XX KW LFA3; CD2; cell signalling; modulation; lymphocyte; T cell;  
 KW memory effector T lymphocyte; psoriatic arthritis;  
 KW rheumatoid arthritis; multiple sclerosis; atopic dermatitis;  
 KW uveitis; inflammatory bowel disease; Crohn's disease;  
 KW ulcerative colitis; cutaneous T cell lymphoma; inhibition; treatment;  
 KW therapy.

XX OS Homo sapiens.

XX PN WO200012113-A2.

XX .PD 09-MAR-2000.

XX PF 31-AUG-1999; 99WO-US20026.

XX PR 31-AUG-1998; 98US-0098456.

XX PA (BIOJ ) BIOGEN INC.

XX PI Magilavy D;

XX DR WPI: 2000-282928/24.

XX DR N-PSDB; 293401.

XX PT Selective modulation of memory effector T lymphocytes by administration  
 of a CD2 binding agent which inhibits the CD2/LFA-3 interaction useful  
 for treating conditions such as inflammatory bowel diseases, psoriatic  
 arthritis

XX PS Disclosure; Page 75-76; 76pp; English.

XX CC Modulation of LFA3/CD2 interaction by administration of a CD2 binding  
 agent inhibits CD2 signalling and T cell proliferation and activation  
 and more particularly modulates the number and/or distribution of  
 memory effector T lymphocytes. The method can be used for treating a  
 condition in a subject where the condition is characterized by memory  
 effector T lymphocytes playing a role in the pathogenesis of the  
 condition such as psoriatic arthritis, rheumatoid arthritis,  
 multiple sclerosis, atopic dermatitis, uveitis, inflammatory bowel  
 disease, Crohn's disease, ulcerative colitis and cutaneous T cell  
 lymphoma and where the method comprises administering to the subject  
 an amount of CD2 binding agent sufficient to modulate the memory  
 effector T lymphocytes. The methods provide inhibition of antigen  
 specific interactions for all antigens present, inhibition of T cell  
 activation, no general immunosuppression, and possibly induction of  
 tolerance.

XX SQ Sequence 347 AA;

Query Match 54.9%; Score 1234; DB 21; Length 347;  
 Best Local Similarity 99.6%; Pred. No. 1e-75; 0; Indels 0; Gaps 0;  
 Matches 227; Conservative 1; Mismatches 0

Qy 1 MDKTHTCPCAPPELLGGPSVFLFPKPKDTLMIARTPEVTCVVVDVSHEDPEVKFNWV 60  
 :|||||  
 Db 120 vdkthtccpapellggpsvflfpkpkdtlmiartpevtcvvvdshedpevkfnwv 179  
 :|||||

Qy 61 DGVEVHNKTKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 120  
 :|||||  
 Db 180 dgvevhnaktpreeqynstyrsvsvltvlhqdwlngkeyckvsnkalpapiektiska 239  
 :|||||

Qy 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 180  
 :|||||  
 Db 240 kgqprepqvylppsrdeltnqvsiltclvkgyfypsdiavewesngqpennyktppvld 299  
 :|||||

Qy 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228  
 :|||||  
 Db 300 sdgsfflyskltdvksrwwqgnvfscsvmhealhnhytqkslsispgk 347  
 :|||||

RESULT 9

W73513 ID W73513 standard; Protein; 388 AA.

XX AC W73513;

XX DT 02-MAR-1999 (first entry)

XX DE Rabbit TGFbetaRII:Fc protein.

XX KW Transforming growth factor-beta receptor; TGF-beta receptor; arthritis;  
 KW fusion protein; fibroproliferative disorder; diabetic neuropathy;  
 KW glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;  
 KW collagen vascular disorder; therapy; rabbit.

XX OS Oryctolagus sp.

XX PN WO9848024-A1.

XX PD 29-OCT-1998.

XX PF 16-APR-1998; 98WO-US07587.

XX PR 18-APR-1997; 97US-0044641.

XX PA (BIOJ ) BIOGEN INC.

XX PI Cate R, Gotwals P, Koteliansky V, Sanicola-Nadel M;

XX WPI: 1998-609994/51.

XX DR N-PSDB; V08998.

XX PT Transforming growth factor-beta receptor fusion protein - used to  
 treat fibroproliferative disorders

XX PS Claim 4; Page 18-19; 70pp; English.

XX CC This sequence is a fusion protein of the rabbit transforming growth  
 factor-beta receptor II (TGF-betaRII) and an antibody Fc region. The  
 encoded protein is an example of a protein of the invention, which are  
 isolated TGF-beta receptor fusion proteins that competitively inhibit  
 binding of TGF-beta to TGF-beta receptor. The fusion protein can be used  
 in a method for lowering the levels of TGF-beta in an individual having  
 arthritis. It can also be used to treat medical conditions such as  
 fibroproliferative disorders. The fibroproliferative disorder is a  
 kidney, intraocular or pulmonary fibrosis, especially selected from  
 diabetic neuropathy, glomerulonephritis, proliferative  
 vitreoretinopathy, or myelofibrosis. The fusion proteins can also be used  
 to treat collagen vascular disorders, including systemic sclerosis,  
 polymyositis, scleroderma, dermatomyositis, and systemic lupus

XX Koteliansky V, Gotwals P, Cate R, Sanicola-Nadel M;  
PI  
XX  
XX WPI; 2000-106083/09.  
DR  
DR N-ESDB; 245251, 245252.  
XX  
XX New fusion protein of a splice variant of transforming growth  
PT factor-beta receptor, for inhibiting the growth factor, e.g. in  
PT treatment of fibrosis -  
XX

PS Disclosure; Page 62-63; 69pp; English.

XX The present sequence represents a splice variant of a rabbit transforming  
 CC growth factor-beta (TGF-beta) type II receptor fused to the Fc portion  
 CC of human IgG1. The fusion protein has higher affinity for TGF-beta than  
 CC fusion proteins comprising the non-variant form of the receptor. The  
 CC fusion protein contains soluble TGF-beta receptor constructs that are  
 CC devoid of a transmembrane region (and are secreted from the cell) but  
 CC retain the ability to bind TGF-beta. The protein competitively inhibits  
 CC binding of TGF-beta to cellular receptors and/or forms an inactive  
 CC complex with TGF-beta. The protein is used to reduce levels of TGF-beta,  
 CC for treatment of arthritic conditions associated with overexpression  
 CC of TGF-beta, especially fibroproliferative diseases, e.g. renal,  
 CC intra-ocular or pulmonary fibrosis; diabetic nephropathy;  
 CC glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;  
 CC collagen vascular disease, e.g. systemic sclerosis, polymyositis,  
 CC scleroderma, dermatomyositis or systemic lupus erythematosus; and  
 CC fibrosis associated with restenosis. It is also used for treating  
 CC wounds, to prevent overproduction of connective tissue and so prevent  
 CC adhesions or scarring, and to prevent post-radiation fibrosis (by  
 CC administration to patients about to undergo radiation therapy).

XX Sequence 388 AA;

Query Match 54.9%; Score 1234; DB 21; Length 388;  
 Best Local Similarity 99.6%; Pred. No. 1.2e-75;  
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKTHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
 :|||||  
 Db 161 vdkthtcpcpapelggpsvflppkpkdtlmisrtpetvcvvdvshedpevkfnwv 220

Qy 61 DGVEVHNKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKKA 120  
 :|||||  
 Db 221 dgvevhnaktpreeqynstyrvvsvltvlhqdwlngkeyckvsnkalpapiektiska 280

Qy 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180  
 :|||||  
 Db 281 kgqprepqvyltppsrdeltnqvsltcclvkgyfypsdiavewesngqpennnykttppvld 340

Qy 181 SDGSFFLYSKLTVDKSRQOQGVNFSCSYVMHEALHNHYTQKSLSLSPGK 228  
 :|||||  
 Db 341 sdgsfflyskltvdksrqoqgvnfvscsvmhealhnhytqkslsispgk 388

RESULT 12  
 Y54064  
 AC Y54064 standard; Protein; 388 AA.  
 XX Y54064;  
 XX  
 XX 27-MAR-2000 (first entry)  
 XX  
 XX Amino acid sequence for TGF-beta type II receptor variant/IgG1 fusion.  
 DE  
 XX Splice variant; human; transforming growth factor-beta; TGF-beta;  
 KW type II receptor; Fc portion; human IgG1; fusion protein; arthritis;  
 KW fibroproliferative disease; renal; intra-ocular; pulmonary; fibrosis;  
 KW diabetic nephropathy; glomerulonephritis; collagen vascular disease;  
 KW proliferative vitreoretinopathy; myelofibrosis; systemic sclerosis;  
 KW polymyositis; scleroderma; dermatomyositis; systemic lupus erythematosus;  
 KW restenosis; wound; connective tissue production; adhesion; scarring;  
 KW post-radiation fibrosis.  
 XX  
 XX Synthetic.  
 OS  
 OS Homo sapiens.  
 XX  
 XX W09965948-A1.  
 XX  
 XX 23-DEC-1999.  
 PD  
 XX 16-JUN-1999; 99WO-US13629.

XX 16-JUN-1998; 98US-0089452.  
 XX (BIOJ ) BIOGEN INC.  
 PA  
 XX Koteliansky V, Gotwals P, Cate R, Sanicola-Nadel M;  
 PI  
 XX WPI; 2000-106083/09.  
 DR  
 XX N-PSDB; 245253, 245254.  
 DR  
 XX New fusion protein of a splice variant of transforming growth  
 PT factor-beta receptor, for inhibiting the growth factor, e.g. in  
 PT treatment of fibrosis  
 PS  
 XX Disclosure; Page 64; 69pp; English.

XX The present sequence represents a splice variant of a human transforming  
 CC growth factor-beta (TGF-beta) type II receptor fused to the Fc portion  
 CC of human IgG1. The fusion protein has higher affinity for TGF-beta than  
 CC fusion proteins comprising the non-variant form of the receptor. The  
 CC fusion protein contains soluble TGF-beta receptor constructs that are  
 CC devoid of a transmembrane region (and are secreted from the cell) but  
 CC retain the ability to bind TGF-beta. The protein competitively inhibits  
 CC binding of TGF-beta to cellular receptors and/or forms an inactive  
 CC complex with TGF-beta. The protein is used to reduce levels of TGF-beta,  
 CC for treatment of arthritic conditions associated with overexpression  
 CC of TGF-beta, especially fibroproliferative diseases, e.g. renal,  
 CC intra-ocular or pulmonary fibrosis; diabetic nephropathy;  
 CC glomerulonephritis; proliferative vitreoretinopathy; myelofibrosis;  
 CC collagen vascular disease, e.g. systemic sclerosis, polymyositis,  
 CC scleroderma, dermatomyositis or systemic lupus erythematosus; and  
 CC fibrosis associated with restenosis. It is also used for treating  
 CC wounds, to prevent overproduction of connective tissue and so prevent  
 CC adhesions or scarring, and to prevent post-radiation fibrosis (by  
 CC administration to patients about to undergo radiation therapy).

XX Sequence 388 AA;

Query Match 54.9%; Score 1234; DB 21; Length 388;  
 Best Local Similarity 99.6%; Pred. No. 1.2e-75;  
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKTHTCPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
 :|||||  
 Db 161 vdkthtcpcpapelggpsvflppkpkdtlmisrtpetvcvvdvshedpevkfnwv 220

Qy 61 DGVEVHNKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKKA 120  
 :|||||  
 Db 221 dgvevhnaktpreeqynstyrvvsvltvlhqdwlngkeyckvsnkalpapiektiska 280

Qy 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180  
 :|||||  
 Db 281 kgqprepqvyltppsrdeltnqvsltcclvkgyfypsdiavewesngqpennnykttppvld 340

Qy 181 SDGSFFLYSKLTVDKSRQOQGVNFSCSYVMHEALHNHYTQKSLSLSPGK 228  
 :|||||  
 Db 341 sdgsfflyskltvdksrqoqgvnfvscsvmhealhnhytqkslsispgk 388

RESULT 13  
 Y70867  
 ID Y70867 standard; Protein; 399 AA.  
 XX Y70867;  
 AC  
 XX  
 XX 31-JUL-2000 (first entry)  
 XX  
 XX Human interferon-beta-1a/mouse IgG2a Fc fusion protein.  
 DE  
 XX Human; interferon-beta-1a; IFN-beta-1a; immunoglobulin; fusion protein;  
 KW angiogenesis; antisclerotic; antiinflammatory; immunosuppressive; cancer;  
 KW cytostatic; virucide; hepatotropic; antiangiogenic; treatment; fibrosis;

KW multiple sclerosis; inflammatory disease; autoimmune disease; hepatitis;  
 KW viral infection; neovascularisation; mouse; IgG2a Fc domain.  
 OS Chimeric - Homo sapiens.  
 OS Chimeric - Mus sp.

XX Key Location/Qualifiers  
 FH 1..166  
 FT /label= natural\_human\_IFN-beta-la  
 FT 167..171  
 FT /label= Enterokinase\_linker  
 FT 172..399  
 FT /label= Mouse\_IgG2a\_Fc\_portion  
 FT /note= "comprises the hinge region, CH2 and CH3  
 FT constant domains of mouse Ig"

PN WO200023472-A2.

XX 27-APR-2000.

XX 15-OCT-1999; 99WO-US24200.

XX 16-OCT-1998; 98US-0104491.

XX 16-FEB-1999; 99US-0120237.

XX (BIOJ ) BIOGEN INC.

XX Whitty A, Runkel L, Brickelmaier M, Hochman P;

XX WPI; 2000-339654/29.

XX N-PSDB; D00167.

XX Fusion proteins comprising interferon-beta-la useful for inhibiting  
 PT angiogenesis -

PS Example 2; Fig 2; 82pp; English.

XX The patent discloses fusion proteins comprising glycosylated  
 CC interferon-beta (IFN-beta) especially IFN-beta-la, linker groups and  
 CC non-IFN-beta proteins, especially an immunoglobulin (Ig) protein.  
 CC The present sequence is a fusion protein that consists  
 CC of mature human IFN-beta-la and mouse IgG2a Fc domain separated by an  
 CC enterokinase linker. The fusion protein is useful for  
 CC inhibiting angiogenesis in a patient. It may also be used to treat  
 CC multiple sclerosis, fibrosis, inflammatory and autoimmune diseases,  
 CC cancers, hepatitis and viral infection characterised by  
 CC neovascularisation.

XX Sequence 399 AA;

Query Match 54.9%; Score 1234; DB 21; Length 399;  
 Best Local Similarity 99.6%; Pred. No. 1.2e-75;  
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPCAPALGPGSVFLFPKPKDTLMSRPEVTCVVVDVSHEDPEVKENNYV 60  
 Db :|||||  
 172 vdkthtcpcpcapellgpgsvflfppkpkdtlmsrptcvvvdvshedpevkfnvyv 231  
 QY 61 DGVEVHNKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISK 120  
 Db :|||||  
 232 dgvevhnaktpreeqynstyrvsvvltvlhqdwlngkeykckvsnkalpapiektiska 291  
 QY 121 KGQPREQVYTLPSRDELKQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180  
 Db :|||||  
 292 kgqpreqvylpsrdeltkngvsltclvkgyfypsdiavewesngqpennnykttppvld 351

QY 181 SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPCK 228  
 Db :|||||

Db 352 sdgsfflyskltvdksrwqqgnvfscsvmhcalhnhytqkslsispck 399

RESULT 14

Y70868

ID Y70868 standard; Protein; 418 AA.

XX Y70868;

XX 31-JUL-2000 (first entry)

XX Human interferon-beta-la G162C-IgG1 Fc direct fusion protein.

XX Human; interferon-beta-la; IFN-beta-la; immunoglobulin; fusion protein;  
 KW angiogenesis; antisclerotic; antiinflammatory; immunosuppressive; cancer;  
 KW cytostatic; virucide; hepatotropic; antiangiogenic; treatment; fibrosis;  
 KW multiple sclerosis; inflammatory disease; autoimmune disease; hepatitis;  
 KW viral infection; neovascularisation; IgG1 Fc domain; VCAM-1; mutant;  
 KW vascular cell adhesion molecule-1.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

XX 1..24

XX /label= VCAM-1\_signal\_sequence

XX /note= "Vascular cell adhesion molecule-1"

XX 25..190

XX /label= natural\_human\_IFN-beta-la

XX Misc-difference 186

XX /note= "Wild type IFN-beta-la Gly at position 162 is

XX substituted by Cys"

XX 192..418

XX /label= Human\_IgG1\_Fc\_portion

XX /note= "comprises hinge, CH2 and CH3 domains of

XX IgG1 heavy chain"

XX WO200023472-A2.

XX 27-APR-2000.

XX 15-OCT-1999; 99WO-US24200.

XX 16-OCT-1998; 98US-0104491.

XX 16-FEB-1999; 99US-0120237.

XX (BIOJ ) BIOGEN INC.

XX Whitty A, Runkel L, Brickelmaier M, Hochman P;

XX WPI; 2000-339654/29.

XX N-PSDB; D00167.

XX Fusion proteins comprising interferon-beta-la useful for inhibiting  
 PT angiogenesis -

PS Example 5; Fig 10; 82pp; English.

XX The patent discloses fusion proteins comprising glycosylated  
 CC interferon-beta (IFN-beta) especially IFN-beta-la, linker groups and  
 CC non-IFN-beta proteins, especially an immunoglobulin (Ig) protein.  
 CC The present sequence is a fusion protein consisting of a vascular  
 CC cell adhesion molecule-1 (VCAM-1) signal sequence, a modified human  
 CC IFN-beta-la and human IgG1 Fc domain, which are directly fused without  
 CC a linker sequence. The fusion protein is useful for  
 CC inhibiting angiogenesis in a patient. It may also be used to treat  
 CC multiple sclerosis, fibrosis, inflammatory and autoimmune diseases,  
 CC cancers, hepatitis and viral infection characterised by  
 CC neovascularisation.

XX Sequence 418 AA;

Query Match 54.9%; Score 1234; DB 21; Length 418;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-75;  
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
Db :|||||  
191 vdkthtcpcpapellggspsvflfppkpkdtlmisrtpevtcvvvdshedpevkfnwv 250  
Qy 61 DGVEVHNAKTKPREEQYNSTYRVSVLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 120  
Db :|||||  
251 dgvevhnaktkpreeqynstyrsvsltlvqhgdwlngkeyckvsnkalpapiektiska 310  
Qy 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 180  
Db :|||||  
311 kgqprepqvyltppsrdeltnqvsltcclvkgyfpsdlavewesngqpennnykttppvld 370  
Qy 181 SDGSFFLYSKLTVDKSRWQQGNVSCSYMHAEALHNHYTQKSLSLSPGK 228  
Db :|||||  
371 sdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispgk 418

RESULT 15  
Y70869  
ID Y70869 standard; Protein; 423 AA.  
XX AC Y70869;  
XX DT 31-JUL-2000 (first entry)  
XX Human interferon-beta-1a G162C/G4S linker/IgG1 Fc fusion protein.  
XX DE Human; interferon-beta-1a; IFN-beta-1a; immunoglobulin; fusion protein;  
KW angiogenesis; antisclerotic; antiinflammatory; immunosuppressive; cancer;  
KW cytostatic; virucide; hepatotropic; antiangiogenic; treatment; fibrosis;  
KW multiple sclerosis; inflammatory disease; autoimmune disease; hepatitis;  
KW viral infection; neovascularisation; IgG1 Fc domain; VCAM-1; mutant;  
KW vascular cell adhesion molecule-1; G4S linker.  
XX OS Homo sapiens.  
OS Synthetic.  
XX FH Key  
XX FT Peptide  
FT 1..24  
FT /label= VCAM-1\_signal\_sequence  
FT /note= "Vascular cell adhesion molecule-1"  
FT Protein  
FT 25..190  
FT /label= natural\_human\_IFN-beta-1a  
FT Misc-difference 186  
FT /note= "Wild type IFN-beta-1a Gly at position 162 is  
FT substituted by Cys"  
FT 191..195  
FT Region  
FT /label= G4S\_linker  
FT 197..423  
FT Region  
FT /label= Human\_IgG1\_Fc\_portion  
FT /note= "comprises hinge, CH2 and CH3 domains of  
FT IgG1 heavy chain"  
XX WO200023472-A2.  
PN 27-APR-2000.  
PD 15-OCT-1999; 99WO-US24200.  
XX 16-OCT-1998; 98US-0104491.  
PR 16-FEB-1999; 99US-0120237.  
XX (BIOJ ) BIOGEN INC.  
XX Whitty A, Runkel L, Brickelmaier M, Hochman P;  
XX WPI; 2000-339654/29.  
DR N-PSDB; D00168.  
XX Fusion proteins comprising interferon-beta-1a useful for inhibiting  
PT angiogenesis -  
XX Example 5; Fig 11; 82pp; English.

XX The patent discloses fusion proteins comprising glycosylated  
CC interferon-beta (IFN-beta) especially IFN-beta-1a, linker groups and  
CC non-IFN-beta proteins, especially an immunoglobulin (Ig) protein.  
CC The present sequence is a fusion protein consisting of a vascular  
CC cell adhesion molecule-1 (VCAM-1) signal sequence, a modified human  
CC IFN-beta-1a which is linked to human IgG1 Fc domain via a G4S  
CC linker sequence. The fusion protein is useful for  
CC inhibiting angiogenesis in a patient. It may also be used to treat  
CC multiple sclerosis, fibrosis, inflammatory and autoimmune diseases,  
CC cancers, hepatitis and viral infection characterised by  
CC neovascularisation.  
XX SQ Sequence 423 AA;  
Query Match 54.9%; Score 1234; DB 21; Length 423;  
Best Local Similarity 99.6%; Pred. No. 1.3e-75;  
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60  
Db :|||||  
196 vdkthtcpcpapellggspsvflfppkpkdtlmisrtpevtcvvvdshedpevkfnwv 255  
Qy 61 DGVEVHNAKTKPREEQYNSTYRVSVLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 120  
Db :|||||  
256 dgvevhnaktkpreeqynstyrsvsltlvqhgdwlngkeyckvsnkalpapiektiska 315  
Qy 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 180  
Db :|||||  
316 kgqprepqvyltppsrdeltnqvsltcclvkgyfpsdlavewesngqpennnykttppvld 375  
Qy 181 SDGSFFLYSKLTVDKSRWQQGNVSCSYMHAEALHNHYTQKSLSLSPGK 228  
Db :|||||  
376 sdgsfflyskltvdksrwqgnvfscsvmhealhnhytqkslsispgk 423

Search completed: March 1, 2001, 09:17:10  
Job time: 228 sec